ChatGPT

# Common Eligibility Criteria in Lung Cancer, Breast Cancer, and Alzheimer's Trials

Clinical trials for lung cancer (carcinoma of the lung), breast cancer, and Alzheimer's disease tend to use recurring inclusion/exclusion criteria across several key domains. Below we summarize the **most frequently occurring criteria** in each domain (demographics, clinical status, diagnosis confirmation, prior treatments, imaging findings, pathology/biomarkers, and laboratory results) for each disease area, based on analyses of trial protocols. We also describe how each criterion's data is represented in hospital electronic medical records (EMRs) – both in **structured form** (e.g. via FHIR resources like *Patient*, *Observation*, *Condition*, *DiagnosticReport*) and in **unstructured documents** (clinical notes, radiology/pathology reports, etc.). Tables are included to quantify how often these criteria appear in trial protocols, and the text elaborates on data representation.

## Lung Cancer Trials (Carcinoma of the Lung)

**Overview:** Lung cancer trials (primarily non-small cell lung cancer, NSCLC) commonly enroll adults with confirmed advanced lung carcinoma and require participants to have good functional status and adequate organ function. Stringent exclusions often relate to brain metastases, prior treatments, comorbid illnesses, and abnormal lab values. **Table 1** outlines common eligibility criteria in lung cancer trials and their frequency. Below, we detail top criteria in each domain and their EMR representations.

**Table 1. Common Lung Cancer Trial Eligibility Criteria (Frequency in Protocols)**

| Domain | Common Eligibility Criteria – Lung Cancer | Prevalence in Trials |
|---|---|---|
| Demographics | Adult age (e.g. **≥18** years; rarely an upper age limit) | Nearly universal (adult trials) [1]; upper age cutoffs used in few trials (e.g. ~19% a decade ago [2]) |
| Clinical Status | **Performance status 0–1** (ECOG scale ≤1) | Majority of trials (≈70–80%) [1] require good PS (exclude ECOG ≥2) |
| | No severe **comorbidities** (e.g. no NYHA class III/IV heart failure, recent MI, uncontrolled illness) | Very common (most trials ~75%+) [3] (e.g. 47% exclude serious cardiac disease [3]) |
| | No active **infections** (HIV, HBV, HCV) or immunosuppression | Common (historically standard; being broadened) [4] |
| | No active **autoimmune disease** (for immunotherapy trials) | Common in immunotherapy trials [5] (standard exclusion to avoid immune-related toxicity) |
| Diagnosis | **Histologically confirmed** lung carcinoma (NSCLC or SCLC) | Universal in trials (pathology report confirming cancer) |

| Domain | Common Eligibility Criteria – Lung Cancer | Prevalence in Trials |
|---|---|---|
| | Advanced **disease stage** (e.g. metastatic Stage IV or unresectable III) | Trial-specific (e.g. all metastatic trials require Stage IV) |
| | Specific **oncogenic driver status** if applicable (e.g. EGFR mutation, ALK fusion for targeted therapy trials) | Many targeted trials; e.g. EGFR mutation required in EGFR inhibitor studies |
| Prior Treatment | **Prior therapy limits** (e.g. "no prior systemic therapy" for 1st-line trials, or ≤1 prior line for 2nd-line trials) | Very common (nearly all trials define allowed prior lines) |
| | **No prior specific drug/class** (to avoid cross-resistance or safety issues) | Common (e.g. exclude prior anti-EGFR in EGFR-targeted trial) |
| | **Washout periods** after previous treatment (e.g. ≥4 weeks since chemo/radiation) | Common (included in most trials) |
| Imaging | No untreated or unstable **brain metastases** (CNS mets must be treated/stable or exclude) | Majority (~80% exclude or conditionally allow brain mets) [6] |
| | **Measurable disease** by RECIST 1.1 (at least one lesion ≥10 mm) | Very common in Phase II/III trials (required for ~70–80% of efficacy trials; qualitative estimate) |
| | No **interstitial lung disease (ILD)** or severe pulmonary fibrosis on scans | Often an exclusion in lung trials [7] (ILD present in ~2–24% of lung CA patients and usually excluded) |
| Pathology/ Biomarker | **Biomarker-defined subsets** required (if applicable): e.g. EGFR mutant, ALK-positive, PD-L1 ≥1% | Many trials (especially targeted therapy trials) focus on biomarker-positive populations |
| | **Histologic subtype** restrictions (e.g. "non-squamous NSCLC only" if drug not suitable for squamous histology) | Common in certain trials (e.g. anti-angiogenic therapy) |
| Laboratory | **Hematologic**: Adequate bone marrow function (ANC ≥1.5×10^9/L, platelets ≥100×10^9/L, Hb ≥9–10 g/dL) | Vast majority (~80–90% of trials) [3] |
| | **Hepatic**: AST/ALT ≤2.5×ULN (≤5× if liver mets), bilirubin ≤1.5×ULN | Vast majority (~80%+) [3] (51% with strict liver criteria [3]) |
| | **Renal**: Creatinine ≤1.5×ULN or CrCl ≥50–60 mL/min | ~90% exclude moderate renal impairment (CrCl <60) [8] |
| | **Viral serologies**: negative for HIV, hepatitis B/C | Common (standard in past trials) [4] |

| Domain | Common Eligibility Criteria – Lung Cancer | Prevalence in Trials |
|---|---|---|
| | **Pregnancy**: negative pregnancy test; no breastfeeding (with contraception required) | Nearly universal in trials with women of childbearing potential |

## Demographic Criteria (Lung Cancer)

- **Age:** Almost all lung cancer trials are restricted to adults (commonly age ≥18). Unlike some older breast trials, explicit upper age limits are now uncommon in lung trials (in line with efforts to include older patients) [2]. Age is recorded in EMRs as structured demographic data in the patient's record (e.g. birth date in the FHIR **Patient** resource), and is often noted in clinical text (e.g. "55-year-old male").
- **Sex:** Lung trials generally include both men and women. Sex is a patient demographic field in the EMR (FHIR Patient.gender). In documentation, sex may be mentioned as part of history (e.g. "female patient with NSCLC"). Pregnancy status is handled via lab tests (see Laboratory criteria).
- **Other:** Trials universally require informed consent capability. Most also implicitly require an adequate support system; however, social factors (marital status, etc.) are usually not explicit criteria in lung cancer trials.

## Clinical Criteria (Lung Cancer)

- **Performance Status (PS):** A *good functional status* is one of the most frequent eligibility criteria. Patients must typically have an Eastern Cooperative Oncology Group (**ECOG**) performance status of 0–1 (fully active or mild symptoms), excluding those who are more debilitated [9]. In one breast cancer cohort, 69% of trials excluded patients with ECOG >1 [1], and lung trials similarly require ECOG ≤1 in the majority. In the EMR, performance status is often recorded as a numeric score in oncology notes (e.g. "ECOG 1"). Structured data capture is possible via an **Observation** with a standard code (LOINC *89247-1* for ECOG score) [10]. This structured entry can be stored in FHIR as an Observation resource (as in the mCODE oncology data model) and retrieved later. Unstructured clinical notes also frequently document PS in free text ("PS 0").
- **Comorbid Conditions:** Trials commonly exclude patients with **significant comorbidities** that could confound results or increase risk. For lung cancer, criteria often exclude those with serious cardiovascular disease (e.g. recent myocardial infarction, uncontrolled angina or NYHA class III/IV heart failure) or stroke, as well as those with poor pulmonary status (other than the cancer) [11] [12]. For example, 47% of breast cancer trials from 2020–22 had strict cardiac-related exclusions [3], and lung trials similarly often require no severe cardiac history. These conditions are tracked in EMRs as diagnosis codes or problem list entries (FHIR **Condition** resources for problems like heart failure or CAD). In clinical documents, comorbidities appear in histories (e.g. "history of congestive heart failure"). Trials may also exclude uncontrolled diabetes, active peptic ulcers, severe COPD, or other significant illnesses per investigator judgment ("no uncontrolled intercurrent illness" is a common phrasing).
- **Prior or Concurrent Malignancies:** Many cancer trials exclude patients with a *recent second malignancy* (other than the cancer under study), typically requiring no other invasive cancer within 2–5 years. In one analysis, 68% of breast trials excluded patients with a prior malignancy history [3]. Lung trials have historically applied similar exclusions. In the EMR, prior cancers appear as **Condition** entries (problem list of cancer diagnoses) or in oncology history notes. These may need manual review to determine eligibility. Structured data (ICD/SNOMED codes for conditions, or **CancerCondition** profile in mCODE) can flag other malignancies, but often a coordinator reviews the narrative history.

- **Infections (HIV/HBV/HCV):** Traditional oncology criteria exclude patients with **active infections** such as HIV or hepatitis B/C due to concerns about drug safety and confounding effects. An analysis noted that many lung trials excluded HIV-positive patients despite modern antivirals making HIV a chronic disease [4]. There is a movement to relax this, but historically a *negative* status for HIV and hepatitis is required. These are confirmed via lab tests (serologies). In the EMR, **Observation** records for HIV/HBV/HCV test results (with LOINC codes for lab tests) provide structured documentation of infection status. Unstructured lab reports or infectious disease notes also contain this information. For trial screening, a documented negative lab result (often scanned or entered as text) is reviewed.
- **Autoimmune Disorders:** Especially in immunotherapy trials, patients with baseline **autoimmune diseases** (e.g. lupus, multiple sclerosis) have often been excluded over concerns that activating the immune system could exacerbate their condition [5]. This exclusion is usually narrowly defined – for example, "no active autoimmune disease requiring systemic immunosuppressive therapy." In practice, the patient's problem list (Condition entries for diseases like rheumatoid arthritis) and specialist notes are reviewed. Structured representation is via **Condition** resources (with SNOMED codes for each autoimmune condition) marked as active. Clinicians also note these in text ("history of autoimmune thyroiditis – exclusionary for immunotherapy trial"). Recent guidelines suggest including patients with stable, well-controlled autoimmune disorders in later-phase trials [13] [14], but early trials often still exclude them.
- **Oxygen Dependence:** Lung cancer trials **frequently exclude** patients who require supplemental oxygen for chronic lung conditions [15]. Requiring home oxygen is viewed as a marker of poor pulmonary function. This criterion ties to comorbidity (often COPD) but is usually stated separately ("exclude patients on continuous oxygen therapy"). In EMRs, long-term oxygen use might be documented in respiratory therapy orders or pulmonology notes rather than a single structured field. Clinicians may note "on 2 L/min $O_2$ at rest" in narrative form. There isn't a dedicated FHIR element for oxygen use status, but it could be captured as an Observation or device use statement. Typically, coordinators determine this from chart notes.
- **General Health and Others:** Trials often require a **life expectancy > 3 or 6 months**, ensuring the patient can likely complete the study. This is a subjective criterion based on physician assessment, not a specific data field – it's documented in text (e.g. "life expectancy approximately 6+ months"). Other general exclusions include pregnancy (discussed under labs) and sometimes neurological status (e.g. no uncontrolled seizures or psychiatric illness that would interfere with compliance). These are evaluated via clinical history and mental status exams documented in notes.

## Diagnosis Criteria (Lung Cancer)

- **Confirmed Lung Carcinoma:** A **histologically or cytologically confirmed diagnosis** of lung cancer is required in essentially all trials [16]. This means a pathology report (from a biopsy or cytology) must verify the patient's tumor type (e.g. "adenocarcinoma of the lung"). In the EMR, the confirmed cancer diagnosis is stored as a **Condition** (often on the problem list, coded as malignant neoplasm of lung) with a verification status of "confirmed" [17]. The supporting pathology report is available as a **DiagnosticReport** resource (category: pathology) with details of the tumor histology [18]. In unstructured form, clinicians often quote "biopsy proven NSCLC" in notes. Trial protocols typically require submission of the pathology report as source documentation.
- **Cancer Stage and Extent:** Trials specify the disease stage appropriate for their study. For lung cancer, most therapeutic trials target **advanced disease** – e.g. metastatic (Stage IV) or unresectable Stage IIIB/IV NSCLC for systemic therapies. Thus, eligibility may read "Stage IV NSCLC" or "recurrent/progressive disease after curative therapy." Staging information in EMRs can be structured (TNM stage recorded as an observation or in a cancer registry module), but

often it's found in oncology notes or tumor board summaries (text). A **Condition** resource can include stage data or references to a **Stage Group** observation. In documentation, one might see "Diagnosis: Stage IV (T2aN2M1) lung adenocarcinoma." Trials also often require *measurable disease* (discussed under Imaging) which implies metastatic lesions that can be measured. Early-stage trials (e.g. adjuvant therapy studies) would conversely require operable, localized disease; those are less common than advanced disease trials.

- **Tumor Subtype and Histology:** Some trials limit enrollment to a particular **histologic subtype** of lung cancer. For example, trials of certain targeted therapies or anti-angiogenic drugs often require *non-squamous NSCLC* (excluding squamous-cell carcinoma due to safety concerns like bleeding risk). This criterion is determined from the pathology report (e.g. "moderately differentiated adenocarcinoma" vs "squamous carcinoma"). In structured data, the cancer histology might be encoded as a SNOMED CT code on the Condition (e.g. adenocarcinoma = SNOMED code) or in a separate pathology Observation. Pathology reports in text explicitly state the subtype. Similarly, small cell lung cancer (SCLC) trials are separate – a trial will specify **"histologically confirmed SCLC"**, excluding NSCLC, and often require extensive vs limited-stage classification.
- **Biomarker-Defined Diagnosis:** An increasing number of lung trials focus on molecularly defined subsets – effectively making a specific biomarker status part of the diagnosis. For example, trials of EGFR inhibitors require **EGFR-mutant NSCLC**, ALK inhibitor trials require **ALK-positive** disease, etc. In such cases, a *molecular diagnostic test* (usually a genetic test on the tumor) must confirm the presence of the mutation/fusion. This is documented in EMRs as a lab or pathology result: often as a **DiagnosticReport** for molecular pathology or **Observation** with the genetic finding (e.g. "EGFR exon 19 deletion detected") using standard codes or genomic variant representations. The FHIR Genomics guidelines capture such results as Observations with complex genetic variant data. In unstructured form, the pathology or molecular report (PDF/text) contains the mutation result. These criteria are common: for instance, EGFR mutations occur in ~10–15% of NSCLC and numerous trials limit enrollment to EGFR-positive patients. PD-L1 expression (a biomarker by immunohistochemistry) is another example – some immunotherapy trials require **PD-L1 tumor proportion score ≥ 1% or 50%**. PD-L1 results are stored either as a pathology report or as an Observation (with a value like "TPS 80%"). Clinicians will note in text "PD-L1 positive (80%)" when considering trial options.
- **Disease Definition:** Trials usually require *no alternate primary cancer* (to ensure the lung tumor is the disease being treated). Also, if the patient has certain tumor behavior (e.g. *no predominant small cell component in a mixed tumor* for NSCLC trials), that might be specified. These nuances are gleaned from pathology reports and are part of the structured diagnosis (e.g. a pathology **DiagnosticReport** would mention any mixed histology).

## Treatment History Criteria (Lung Cancer)

- **Prior Lines of Therapy:** Eligibility often limits the amount of **prior treatment** a patient has received for their lung cancer, to define a relatively uniform population. For first-line therapy trials in advanced NSCLC, typically **no prior systemic therapy** for advanced disease is allowed (though prior adjuvant therapy might be allowed if completed >6–12 months before recurrence). For second-line trials, they may require exactly one prior line of chemotherapy or immunotherapy. Thus, criteria like "must have progressed after exactly one prior platinum-based regimen" or "no more than two prior systemic regimens for metastatic disease" are common. Nearly all trials specify such prior treatment limits. In the EMR, a patient's treatment history is recorded through **MedicationAdministration/Statement** resources for drugs given, or **Procedure** resources for surgeries and radiotherapy. Oncologists also summarize prior treatments in notes ("received carboplatin-pemetrexed first line, progressed after 8 months").

Determining eligibility often involves manually reviewing these records or structured oncology module data for prior lines.

- **Previous Drug Exposures:** Trials frequently exclude patients who have taken a drug similar to the investigational agent. For example, a trial of a new EGFR inhibitor will say **"no prior EGFR tyrosine kinase inhibitor"**. Similarly, an immunotherapy trial might exclude those who previously received a PD-1/PD-L1 inhibitor. This avoids confounding by resistance or prior exposure. These details rely on medication records – in FHIR, **MedicationStatement** (list of past meds) or **Procedure** (if a therapy is considered a procedure) can hold this. Unstructured sources include oncology notes or pharmacy records; e.g. "Patient was on osimertinib – exclusion for trial of another EGFR TKI."
- **Washout Periods:** Almost all trials impose **washout periods** to ensure prior treatments have cleared and won't interact. For example, "at least 4 weeks since last chemotherapy or major surgery" or "≥2 weeks since radiation therapy (palliative)" are standard. The criterion might be stricter for certain drugs (e.g. 6 weeks if a drug has long half-life or if prior therapy was an antibody). This information is derived from treatment dates in the EMR. Structured data: each **MedicationAdministration** or **Procedure** has a date; one could compute the interval. In practice, trial coordinators check the last treatment date documented in progress notes or chemo administration records.
- **Concurrent Therapies:** Trials typically prohibit other ongoing cancer treatments during the study. So criteria include no concurrent chemotherapy, hormonal therapy (unless specified), or other investigational agents. This is checked by ensuring the patient has discontinued prior regimens (EMR medication lists should not show active chemo orders). It's also often explicitly documented in a consent note when transitioning to trial. FHIR **MedicationRequest** could indicate if an anti-cancer med is currently active. In text, one might see "Patient stopped erlotinib 1 month ago to enroll in trial."
- **Prior Surgery/Radiation:** Some trials require that any major surgery or radiation therapy be completed and the patient **recovered**. For example, "recovered from toxicity of prior therapy to ≤Grade 1" (meaning any side effects are mild) is common. Also, "at least 4 weeks post-surgery and wounds healed." These are clinical judgments documented in notes (e.g. "post-operative recovery complete"). There might not be a discrete EMR field for "recovered," but the absence of active wound care orders or an exam note stating good healing serves as evidence. Past surgeries are stored as **Procedure** records in structured data.
- **Prior Malignancy:** (As noted under clinical comorbidities) if a patient had another cancer recently, they might be excluded unless disease-free for a certain period. The presence of another cancer and its treatment timeline would be reviewed via the problem list and historical notes.

## Imaging Criteria (Lung Cancer)

- **Brain Metastases:** The presence of **brain metastases** is a critical eligibility factor. Historically, most lung cancer trials *excluded patients with active or untreated brain metastases*, due to concerns about poor prognosis and drug penetration. Patients with previously treated brain mets (surgically resected or radiated) who are neurologically stable and off steroids are often allowed ("stable CNS metastases allowed"). An analysis across cancers (2012–2022) found trials are becoming more inclusive: the proportion of trials **allowing** brain metastasis rose from 11.5% to 17.3%, while those with conditional exclusion (i.e. allowed if treated/stable) remain ~75% [6]. In practice, current lung trials typically require an MRI brain scan at screening; if it shows untreated or progressing mets, the patient is excluded. EMR data for brain mets comes from radiology reports (text) and problem lists (some patients have a diagnosis entry like "brain metastasis" (Condition)). There is no binary structured flag for "has brain mets"; one must read the **DiagnosticReport** of a brain MRI. FHIR can represent the imaging study (as an

**ImagingStudy** and a radiology **DiagnosticReport**), but the interpretation "brain metastases present" is usually in the narrative impression [19]. Clinicians will note "MRI brain: no metastases" or "3 brain mets treated with SRS" in their notes. Trial criteria around brain mets are thus assessed via unstructured imaging results and oncology history.

- **Measurable Disease:** Many interventional trials (especially Phase II/III with tumor response endpoints) require patients to have **measurable disease by RECIST 1.1 criteria**. This means at least one lesion ≥10 mm on CT (or ≥20 mm on chest X-ray, though CT is standard) that can be followed for response [20]. If a patient only has non-measurable disease (e.g. bone lesions without a soft-tissue component, or pleural effusion only), they might be ineligible. The frequency is high – essentially all trials looking at tumor shrinkage require measurable disease. In the EMR, tumor measurements are found in radiology reports (text); some oncology systems allow radiologists to label target lesions and record diameters (structured in tumor registry or via **Observation** entries for each lesion size). However, eligibility determination usually involves a radiologist/clinician reviewing scans. They may document "measurable disease present" in an eligibility note. FHIR could capture a lesion size as an Observation, but in practice one relies on narrative reports like "CT scan shows a 2.5 cm lung mass (measurable)".
- **No Extensive Disease that Precludes Assessment:** Some trials exclude bulky disease or certain patterns. For example, an uncontrolled malignant pleural effusion can make response assessment tricky. While not always explicit, a trial might require effusions to be tapped if large, or exclude patients with leptomeningeal carcinomatosis (since it indicates very advanced disease and poor prognosis). These criteria, when present, are qualitative and verified by imaging reports and clinical exam. They appear in text (e.g. "massive ascites" noted by radiologist might be flagged by investigator as exclusionary if protocol says no significant effusions).
- **Pulmonary Imaging Findings (ILD):** As noted, the presence of **Interstitial Lung Disease (ILD)** is often an exclusion criterion [7]. ILD (chronic lung fibrosis or pneumonitis) can both limit a patient's lung reserve and pose a risk if the trial drug has pulmonary toxicity. Many lung cancer patients have smoking-related lung changes; only clinically significant ILD (e.g. idiopathic pulmonary fibrosis with symptoms) typically triggers exclusion [21]. Radiology CT reports are used to identify ILD. Structured data for ILD might be a radiology result (often just text: "bibasilar reticulation consistent with fibrosis"). If ILD is known, it may be on the problem list (Condition: interstitial pulmonary disease). Trial protocols sometimes specifically mandate screening CT review for ILD. In FHIR, an **Observation** could record an imaging finding (e.g. presence of fibrosis), but this is not commonly automated – it's manual radiologist judgment. Thus, the investigator reads the report (unstructured) to decide eligibility.
- **Other Imaging Exclusions:** Some trials exclude patients with certain **metastatic locations** that pose high risk. For instance, tumors encasing major blood vessels (risk of hemorrhage) were excluded in early anti-angiogenic trials. Also, if a trial requires repeated biopsies, patients must have an accessible lesion on imaging (so imaging is used to confirm a safe biopsy target). These criteria are less universal but when present, are assessed via imaging review. For example, exclusion might say "no centrally located tumors invading blood vessels" – requiring a CT scan review by a radiologist. There's no discrete EMR field for "tumor invasion of vessel" – it's in the radiology report narrative. An **ImagingStudy** in FHIR would reference series/images, but the key info is again in the radiologist's text conclusion.
- **Imaging for Staging:** If the trial is in non-metastatic setting, eligibility might require imaging confirmation that disease is not metastatic. E.g., a trial for stage III lung cancer may require a negative brain MRI and abdominal CT to ensure no distant mets before enrollment. Those are procedural requirements rather than criteria ("must undergo staging scans"). The *results* of those scans feed into inclusion ("confirmed stage III, no distant metastasis"). This overlaps with diagnosis/staging criteria and uses the same EMR data (radiology reports and staging documentation).

- **Central Nervous System Performance:** For trials involving brain (e.g. brain mets trials), imaging is used to assess that metastases are not causing significant symptoms or are stable in size. This merges with clinical neurological exam findings. In documentation, one might see "Brain MRI stable – ok per criteria" in a clinic note.

## Pathology/Biomarker Criteria (Lung Cancer)

- **Molecular Biomarkers:** Modern lung cancer trials are often biomarker-driven. As such, **genomic or molecular biomarkers** can be explicit eligibility requirements. For example:
- *EGFR mutation status*: Trials of EGFR inhibitors (erlotinib, osimertinib, etc.) historically required an EGFR-sensitizing mutation. If a trial is for EGFR-mutant patients, then having an EGFR mutation (e.g. exon 19 deletion or L858R) is a top inclusion criterion. Similarly, ALK inhibitor trials require ALK gene rearrangement, ROS1 trials require ROS1 fusion, KRAS G12C inhibitor trials require that specific KRAS mutation, etc. Each of these is determined by a lab test (typically next-generation sequencing panel or PCR). In FHIR, such results are modeled as **Observation** resources (with genetic variant details) or as part of a **DiagnosticReport** (molecular pathology report) [19] . For example, a FHIR Observation might have code "EGFR gene mutation analysis" and a value like "EGFR exon19 deletion present (positive)" coded with a variant ID. In unstructured form, the pathology report text will say "EGFR mutation detected." The trial site will confirm the presence of the mutation in the patient's records (often requiring the lab report as documentation). This criterion is common – numerous trials have been subdivided by these driver mutations since ~2010s.
- *PD-L1 expression*: Some immunotherapy trials (especially first-line NSCLC with single-agent PD-1 inhibitors) required **PD-L1 ≥50%** on tumor cells. PD-L1 is measured via immunohistochemistry and reported as a Tumor Proportion Score (TPS). The pathology report might be a PDF or an entry, e.g., "PD-L1 IHC: 80% TPS (positive)." In structured terms, one could have an Observation with LOINC code for PD-L1 expression and value = 80% (and an interpretation "Positive"). In practice, this might not be fully structured in the EMR; it could be in a scanned report or text note. Clinicians note it in their narrative ("PD-L1 high, qualifies for pembrolizumab trial"). Many immunotherapy trials after 2016 stratified by PD-L1, though some allowed all-comers.
- *Other biomarkers*: Emerging biomarkers like **MET exon 14 skipping**, **RET fusion**, or **BRAF V600E** – each relevant to specific targeted drugs – can be criteria in trials for those subpopulations. Individually these occur in a small percentage of lung cancers, but collectively a significant portion of trials now specify at least one biomarker. The data representation is similar: a specialized lab result, often stored as part of genomic test output. If the hospital uses an integrated genomics system, these might be discrete (e.g. in an **Observation** with a variant ID). Otherwise, they are found in PDF reports attached to the chart.
- **Protein/Biochemical Markers:** Less common in lung than genomic, but e.g. **TPS≥1%** PD-L1 or **TMB (Tumor Mutational Burden)** high could be criteria for some trials. TMB is a number from genomic testing, sometimes included in reports ("TMB = 12 muts/Mb"). That could be an Observation with a numeric value. Generally, any biomarker criterion means that piece of data must be present in the EMR (structured or not) – if not, the patient might need testing to generate it. Trials often screen patients by ordering the required biomarker test if not already done.
- **Histopathologic Features:** Other pathology-based criteria can include **tumor differentiation** (rarely used), or requiring **"measurable disease in pathology"** for certain trial designs (e.g. a trial of a diagnostic agent might require enough tumor tissue). For lung, a key pathology criterion was historically **"no squamous histology for certain drugs"** (as mentioned). That is derived from the pathology report (diagnosis line). In EMR, that's the cancer histology field. Another example: trials of adjuvant therapy in completely resected patients require *pathologic*

*stage II or III* – which is a pathology-based staging after surgery (found in surgical pathology reports and recorded as part of the cancer case summary).

- **Pathology Report Availability:** Practically, a criterion might state that a representative tumor sample (slides or blocks) must be available for central testing. This isn't a patient characteristic per se, but it's eligibility-related. It would be documented as "tissue available: yes" by trial staff after checking the pathology lab. In FHIR, one might represent the sample as a **Specimen** resource linked to a DiagnosticReport, but usually it's managed offline or via notes.
- **Representation in EMR:** Pathology and biomarker criteria rely on **DiagnosticReport (Pathology)** entries in FHIR, which encapsulate the findings of tissue exams [18]. Within those, specific findings like cancer type or biomarker results can be coded Observations. For instance, the **DiagnosticReport** may have a conclusion like "Non-small cell lung carcinoma, adenocarcinoma subtype, EGFR mutation positive" and reference Observation components for EGFR mutation details. In unstructured documents, the pathology report is a textual document often scanned or interfaced from the lab. Physicians also summarize key pathology in their notes (e.g. "EGFR+: yes, ALK: no, PD-L1: 70%"). These serve as quick references when determining trial fit.

## Laboratory Criteria (Lung Cancer)

Almost all interventional oncology trials include a set of **laboratory value thresholds** to ensure patients have adequate organ function to safely receive the study treatment. Lung cancer trials are no exception – lab-based inclusion criteria are extremely common. These typically cover hematologic, hepatic, renal, and sometimes metabolic function, as well as pregnancy status for women. According to one analysis, 84% of recent breast cancer trials had criteria related to bone marrow function, 51% had strict liver function cut-offs, and 42% had strict renal function cut-offs [3]. Lung trials use similar lab criteria (indeed, a *review of ~90 lung trials found 90% excluded patients with creatinine clearance <60 mL/min* [8]). Key lab criteria and their representation are as follows:

- **Hematological Function:** Patients must have adequate bone marrow reserves. Common cut-offs are:
- **Absolute Neutrophil Count (ANC) ≥ 1.5 × 10^9/L** (1500/μL)
- **Platelet count ≥ 100 × 10^9/L** (100,000/μL)
- **Hemoglobin ≥ 9 or 10 g/dL** (often allowing transfusion to reach this level before enrollment)

These criteria appear in nearly every trial (one study showed 84% of trials had *some* bone marrow criteria [3]). In the EMR, lab results are stored as structured data points: each CBC (complete blood count) produces Observations for WBC, neutrophils, platelets, hemoglobin, etc. In FHIR, each lab result is an **Observation** with a code (e.g. LOINC code for neutrophils) and a value and unit [22]. For example, ANC might be represented by LOINC code 26499-4 "Leukocytes, neutrophils/100 leukocytes" in combination with WBC to calculate absolute count. More commonly, absolute neutrophil count is directly reported by the lab. Trial monitors will check the latest lab report – either viewed in the EHR or on a printout. In text, eligibility checklists might note "ANC 1.8, ok" or if below threshold "ANC too low – patient ineligible." The **DiagnosticReport** resource for a lab panel (CBC panel) can group these Observations [22]. Many sites have alerts if a patient doesn't meet lab criteria. Unstructured lab reports (often a table of values) are also reviewed. - **Liver Function:** Criteria ensure no severe hepatic impairment: - **Bilirubin ≤ 1.5 × upper limit of normal (ULN)** (if no liver metastases; sometimes ≤ 3 × ULN if Gilbert's syndrome or similar benign elevation is allowed) - **AST (SGOT) and ALT (SGPT) ≤ 2.5 × ULN** (or ≤5 × ULN if liver metastases are present, since mild elevation is expected with liver tumors) - Sometimes **alkaline phosphatase ≤ 2–5 × ULN** (particularly if relevant to liver/bone mets).

Strict liver enzyme criteria were present in about 51% of recent breast cancer trials [3] and are standard in lung trials too. Lab Observations in the EMR for these are typically part of a **comprehensive metabolic panel** (CMP). FHIR Observation resources would carry codes like LOINC 14629-0 for AST, 1742-6 for ALT, 1975-2 for bilirubin, each with values and reference ranges. The EMR can automatically flag if a value is above the threshold. Representations: in structured form, these values come with interpretation flags (high/normal). A FHIR **Observation.interpretation** might indicate if it's high [23]. In trial documentation, one might see "AST 1.2 × ULN, eligible" or if too high "bilirubin 3.0 (ULN 1.2) – excludes patient." The **DiagnosticReport** for the liver panel will include these Observations, which can be referenced for decision support. - **Renal Function:** Adequate renal function is required, usually defined as: - **Serum creatinine ≤ 1.5 × ULN** *or* calculated **creatinine clearance (CrCl) ≥ 50–60 mL/min** (using Cockcroft-Gault or similar formula).

In fact, an analysis of lung cancer trials found ~90% excluded patients below a CrCl threshold of ~60 mL/min [8]. This is despite many lung cancer patients having some chronic kidney disease; hence stakeholders have argued some cutoffs are unnecessarily high [8]. EMR lab data includes serum creatinine (LOINC 2160-0) as an Observation. Some EHRs calculate CrCl or eGFR automatically (Observation for eGFR could be LOINC 2164-2, "eGFR non-African American", etc., with value in mL/min/1.73m$^2$). If not, the trial coordinator may use a calculator with data from EMR (age, weight, creatinine). FHIR could support on-the-fly calculation via a Decision Support rule or have an Observation for CrCl if the system stores it. In notes, one often reads "Cr 1.0, CrCl ~80 mL/min – meets criteria" or "CrCl 45 mL/min, ineligible (cutoff 50)." - **Metabolic and Other Labs:** Trials may include additional lab exclusions: - **Blood glucose**: Not usually in oncology eligibility unless extreme. (However, poorly controlled diabetes might be excluded under general health, and specific trials like those with high-dose steroids could exclude uncontrolled diabetics). - **Electrolytes**: Generally must be within reasonable range (significant electrolyte imbalances would be corrected before enrollment). - **Coagulation**: Some trials, especially if a drug may cause bleeding, require **INR** and **aPTT** within normal range if not on anticoagulation. If patients are on warfarin, they might be excluded or asked to switch to heparin. These lab results (INR, etc.) are Observations too. - **Cardiac enzymes**: Rarely, if a drug is known to cause cardiac issues, they may check **LVEF** (see below under breast cancer for HER2 therapies) or baseline **troponin**. - **Thyroid function**: For drugs like immune checkpoint inhibitors (which can cause thyroiditis) or others affecting thyroid, baseline **TSH/T4** might be required normal. It's not a universal criterion but can appear.

Generally, any such criteria correspond to specific Observations in lab results. They would be reviewed similarly via lab reports in the EMR. - **Pregnancy Test:** All trials involving women of childbearing potential require a **negative pregnancy test** (usually serum or urine β-HCG) before enrollment, and agreement to use effective contraception. This falls under lab domain (an HCG test result) and demographic (reproductive status). In practice, this is nearly universal: an HCG lab is done within a few days before starting the trial drug. The result is an Observation in EMR (LOINC 2106-3, "HCG beta subunit [Presence] in Serum or Urine"). A negative result is documented; in FHIR it could be an Observation with value "negative" (and often with Observation.code for pregnancy test and Observation.valueBoolean = false or valueCodeableConcept = "Negative"). Unstructured, it might be in a lab report or noted "preg test negative on mm/dd/yyyy" in a research coordinator's note. Additionally, trials exclude **breastfeeding** women – this is typically determined by patient self-report (documented in history and often as an item on a checklist). It's not a standard structured field in EMRs (though lactation status might be noted in obstetric records). Instead, a note "not breastfeeding" suffices. - **Lab Data in EMRs:** Laboratory results in EHRs are highly structured. Each test comes in as a coded result, and FHIR allows grouping them by a lab **DiagnosticReport** (which represents the lab report containing individual test Observations) [22]. For example, a **DiagnosticReport** for "Basic Metabolic Panel" would reference Observations for creatinine, glucose, etc. When assessing eligibility, clinicians look at the latest lab report. Many sites have flowsheets or dashboards highlighting if a value is out of range for trials. In the FHIR context, one could query for the patient's labs via an API (e.g. search Observations by codes for

ANC, etc.) to programmatically check criteria. Indeed, **Observation.category = laboratory** can filter lab results [24]. But in reality, much of this checking is manual or via local clinic research software. Regardless, the structured nature of lab data makes it one of the easier eligibility factors to auto-check. - **Unstructured Lab Data:** Sometimes, especially historically, lab results might be scanned or faxed if outside the system. But nowadays most come electronically. If needed, staff will transcribe outside labs into the record. During trial screening, it's common to have a worksheet listing each required lab and the patient's value (filled from the EMR). If any value doesn't meet the criterion, the patient may be retested or deemed ineligible.

## Breast Cancer Trials

**Overview:** Breast cancer trials share many similar eligibility themes with lung cancer (adult patients, good performance status, adequate organ function) but also have unique criteria related to breast cancer subtypes (hormone receptor and HER2 status), sex (mostly female patients), and cancer stage (early vs metastatic). Recent analyses of breast cancer trial criteria highlight concerns that many trials exclude older patients, those with comorbidities, and those with performance status limitations [25] [26]. **Table 2** summarizes common criteria in breast cancer trials. We then detail each domain's top criteria and how they are recorded in EMRs.

**Table 2. Common Breast Cancer Trial Eligibility Criteria (Frequency in Protocols)**

| Domain | Common Eligibility Criteria – Breast Cancer | Prevalence in Trials |
|---|---|---|
| Demographics | Adult women (often **female sex** only; male breast CA rare) | Nearly all breast CA trials (male inclusion uncommon) |
| | Age criteria (≥18; occasionally **upper age limit** e.g. ≤70) | Upper age limit present in 39% of trials [1] (increased from 19% historically) |
| | Postmenopausal or premenopausal status (trial-specific) | Many endocrine therapy trials specify menopausal status |
| Clinical Status | ECOG **performance status ≤1** required (exclude PS >1) | ~69% of trials require PS 0–1 [1] |
| | **Comorbidities:** no severe cardiac disease, etc. | Very frequent (77% had strict comorbidity exclusions [1]; e.g. 47% cardiac, 42% renal strict limits [3]) |
| | **LVEF ≥50%** if therapy is cardiotoxic (HER2-targeted, anthracyclines) | Standard in HER2+ trials (all require baseline LVEF normal) |
| | No active infection (HIV, HBV, HCV) or pregnancy | Common (standard safety exclusion) |
| Diagnosis | Histologically confirmed **breast carcinoma** (invasive or DCIS per trial) | Universal (pathology confirmation required) |
| | Cancer **subtype** requirements: e.g. **HER2-positive**, or **ER+ and HER2-**, or **Triple-negative** | Very common – trials are often subtype-specific (e.g. all HER2 trials require HER2 overexpression) |

| Domain | Common Eligibility Criteria – Breast Cancer | Prevalence in Trials |
|---|---|---|
| | Disease stage: early (localized) or metastatic as per trial's focus | Specified in all trials (e.g. "metastatic breast cancer" for advanced trials) |
| Prior Treatment | Prior therapy limits (e.g. "≤1 prior chemo for metastatic setting") | Common (especially in metastatic trials) |
| | No prior use of drug class under study (e.g. no prior CDK4/6 inhibitor if testing a new one) | Common in many trials |
| | Prior adjuvant therapy allowed or not (depending on trial phase) | Specified (varies by trial design) |
| Imaging | No uncontrolled **brain metastases** (allowed if treated & stable) | Common (brain mets often excluded unless stable; ~43% trials allowed treated brain mets [6]) |
| | **Measurable disease** if required by endpoints (for metastatic trials) | Very common in metastatic trials (most phase II/III need measurable disease) |
| | Baseline **LVEF imaging** (echocardiogram) normal if required | Required in all HER2+ and anthracycline trials (to meet LVEF ≥50% criterion) |
| Pathology/ Biomarker | **ER/PR and HER2 status** specified (to define subtype) | Nearly all trials: e.g. HER2+ trials (15–20% of breast CA) [27], hormone receptor positive trials, etc. |
| | **BRCA1/2 mutation** status if germline therapy trial (PARP inhibitors) | Many PARP inhibitor trials require BRCA mutation (small subset overall) |
| | **PD-L1 status** for immunotherapy in TNBC (e.g. PD-L1 positive TNBC) | Some trials (e.g. atezolizumab in PD-L1+ TNBC) require this |
| Laboratory | Hematologic: ANC ≥1.5, Platelets ≥100, Hb ≥9–10 g/dL | ~84% trials include (strict or moderate) [3] |
| | Hepatic: AST/ALT ≤2.5×ULN (≤5× if liver mets), bilirubin ≤1.5×ULN | Most trials (51% strict limits) [3] |
| | Renal: Creatinine ≤1.5×ULN or CrCl ≥50–60 mL/min | Most trials (42% strict limits) [3] |
| | Others: LVEF ≥50% (cardiac ultrasound), negative pregnancy test | LVEF required in ~100% of HER2 drug trials; Pregnancy test universal for women of childbearing age |

### Demographic Criteria (Breast Cancer)

- **Sex and Eligibility:** The vast majority of breast cancer trials enroll **female patients** only, since breast cancer predominantly affects women. Male breast cancer is rare (~<1% of cases) and often excluded by trial protocols or capped at a very low number. Thus, being a female is

effectively an inclusion criterion in most cases (some modern trials do allow males, but recruitment of males is very low). In the EMR, patient sex is a core demographic field (FHIR Patient.gender). If a trial does allow men, typically the pathology must confirm breast carcinoma in a male patient, and sometimes separate consent or considerations are noted due to small numbers. In text, trial listings often explicitly state "Female patients age ≥X (male patients may be allowed in some trials of breast cancer)."

- **Age:** Breast cancer trials are for adults, usually **age 18 or older**. Unlike pediatric cancers, breast cancer in minors is essentially nonexistent, so this is given. More notably, some breast trials historically set an **upper age limit** (like "≤65" or "≤70") under assumptions about fitness, although this practice has been criticized. A recent study found 39% of breast trials (2020–22) had an upper age cutoff, up from 19% a decade prior [1], potentially limiting older patient enrollment. For example, a trial might say "Age 18–70 years." In EMR, age is calculated from birth date (Patient.birthDate). If an upper age is in place, it's an objective check – e.g., a 72-year-old would be automatically ineligible if cutoff is 70. Investigators increasingly avoid upper age exclusions to improve representativeness [28]. In trial documents, if an older patient is considered, they might need an exemption or not be allowed per criteria.
- **Menopausal Status:** Some breast cancer trials, especially those involving hormonal therapies, specify **menopausal status**. For instance, a trial for a new aromatase inhibitor might require participants to be postmenopausal (or if premenopausal, to undergo ovarian suppression). Conversely, a trial might focus on premenopausal women specifically. Menopausal status is not a straightforward structured field in general EHR, but it can be inferred from age, menstrual history, or lab tests (FSH levels). In notes it might say "postmenopausal per patient (no menses for 2 years)" or "premenopausal – on LHRH agonist for ovarian suppression." If required, clinicians document it and sometimes lab evidence (like elevated FSH) is used. FHIR could represent it as an Observation (e.g. LOINC code for menopausal status exists in some vocabularies) or simply as an attribute on a research form. This criterion is common: about half of hormone-receptor-positive breast CA trials stratify or specify menopause in some way.
- **Consent and Caregiver:** Breast cancer patients usually do not require caregiver consent as in Alzheimer's, since they retain decisional capacity. However, if a trial requires heavy visit commitment, sometimes they note the patient must be willing and able to comply with study visits (implying adequate support). This is rarely explicit beyond the generic "able to provide informed consent." For older patients, having a caregiver is beneficial but not mandated. Demographically, trials might also collect race/ethnicity for analysis, but they don't typically include it as a criterion (except in certain disparity studies).
- **Reproductive Status:** Women of childbearing potential in breast cancer trials must use effective contraception. This is an almost universal criterion (given systemic therapies can harm a fetus). It's usually listed under both demographic and lab (pregnancy test) criteria. EMRs may have a field for pregnancy status or number of children, but usually it's addressed via counseling and documentation in consents ("patient is not planning pregnancy and will use contraception").
- **Representation:** Demographics like age and sex are captured in **Patient** resource in FHIR (birthDate, gender). Menopausal status or pregnancy intention is less standardized; it might appear in **QuestionnaireResponse** or custom Observation. In clinical practice, these are handled via interview and forms. For trial screening, a coordinator often verifies age from the chart, confirms sex, and ensures the patient falls in the needed category (e.g. if trial requires postmenopausal, they verify that in the history).

## Clinical Criteria (Breast Cancer)

- **Performance Status:** As with lung trials, breast cancer trials almost always require a good **ECOG performance status (PS)**. Typically ECOG 0–1 is required; occasionally a more lenient trial allows ECOG 2, but 0–1 is the norm (69% of trials explicitly excluded ECOG >1 in one cohort [1]). This

criterion ensures patients are physically well enough to handle trial therapy. EMR handling is similar as described before: PS is noted in oncology clinic notes and can be stored as an Observation with ECOG score (0,1 etc.). For example, the oncologist might have a flowsheet where they click ECOG 1 at each visit. That could map to a FHIR Observation (LOINC 89247-1, etc.). If a patient's PS is 2 (e.g. cannot work but up and about >50% of waking hours), they would be excluded. This is often one of the first checks. In text, one might see "PS 2 – not eligible for trial X that requires 0–1."

- **Comorbidities and Organ Function:** Breast cancer predominantly affects women in midlife and older, many of whom have other health issues. Trials often **exclude significant comorbid conditions** to avoid confounding safety. A comprehensive study found 77% of breast trials had at least one strict exclusion related to comorbidities [1]. Common ones:
- **Cardiovascular:** Since certain breast cancer therapies (like anthracyclines or HER2-targeted drugs) stress the heart, trials exclude those with serious cardiac histories. For example, **no recent myocardial infarction, unstable angina, uncontrolled arrhythmia, or Class III/IV heart failure.** Also, a baseline **Left Ventricular Ejection Fraction (LVEF) below normal** is a reason for exclusion (more on LVEF below). The EMR problem list and cardiology notes provide structured data on heart conditions (ICD codes for HF, etc.). LVEF is usually from an echocardiogram report (DiagnosticReport in cardiology). Trials explicitly require LVEF measurement (see Imaging). For structured data, LVEF might be an Observation (LOINC 33878-0 "Left ventricular ejection fraction by Echo"). Many EHRs have cardiology modules capturing this. If a patient has, say, a history of congestive heart failure with EF 40%, they would fail eligibility for a trial requiring normal EF. Clinicians have to make sure any required cardiac clearance is done.
- **Hepatic:** Aside from lab LFT values, actual liver diseases (cirrhosis, active hepatitis) can exclude someone. E.g. "No severe liver disease" might be stated. These are less common than just using labs. But if a patient has known cirrhosis, even if labs are okay, an investigator might deem them ineligible for safety. EMR diagnosis (Condition: cirrhosis) would be relevant.
- **Renal:** Similarly, aside from creatinine, any *clinically significant renal disease (e.g. on dialysis)* would exclude a patient, though such scenarios are rare in active breast cancer patients because dialysis patients might not pursue trials. It might be implicitly handled by lab criteria.
- **Infection/HIV:** Breast trials historically also exclude HIV+/active hepatitis patients, though if someone is well-controlled on HIV meds, some trials may allow them now. Officially, many protocols still list "no active uncontrolled infection" which covers HIV with detectable viral load or hepatitis with positive viral load. This overlaps with lab and comorbidity. In EMR, an active HIV infection is captured by a Condition (B20 or similar code) and labs (HIV viral load). If a patient is HIV-positive but undetectable on ART, investigators often have to check protocol specifics or request sponsor approval.
- **Other Cancer History:** As with lung, breast trials often exclude any other malignancy in the past 3–5 years (except low-risk like non-melanoma skin cancer or in-situ carcinoma). This is to ensure any outcomes can be attributed to breast cancer treatment and not another cancer. EMR problem list listing another cancer (like colon cancer 2 years ago) would trigger this exclusion. Structured data wise, you'd check if any Condition has a cancer code (and see its dates).
- **General Health:** Trials often require that patients be in generally stable health aside from cancer. Phrasing can be broad: "No uncontrolled intercurrent illness including… (list of examples)." This catch-all includes severe psychiatric illness, active substance abuse, severe infections, etc. These are assessed via the physician's judgment documented in notes ("The patient has uncontrolled depression – might not comply with protocol"). It's not strictly structured; rather, it's a medical judgment call often.
- **LVEF (Cardiac Ejection Fraction):** A critical **criterion in HER2-positive breast cancer trials** (and any trial with potentially cardiotoxic agents) is a normal left ventricular ejection fraction. HER2-targeted monoclonal antibodies (e.g. trastuzumab) can cause cardiomyopathy, so trials

uniformly require **LVEF ≥ 50%** (or sometimes ≥55% depending on lab normal). For example, a trial of a new HER2 drug will require a baseline echocardiogram or MUGA scan showing EF in normal range. One analysis showed 47% of trials had strict cardiac criteria [3], which largely reflects LVEF requirements. In the EMR, an echo report provides LVEF (often given as a percentage). This is structured in the cardiology report (some EHRs have it discretely, others one reads the report text). FHIR would use a **DiagnosticReport** for the echo and an **Observation** for LVEF measurement. For instance, Observation.code = "Ejection fraction" and Observation.value = 60%. Unstructured, the echo report might say "LVEF ~60% by Simpson's method." Trials usually require a copy of that report. The research coordinator ensures an echo is done within a certain window pre-trial (say within 6 weeks). If LVEF is 45%, the patient is excluded. If it's borderline (50-55%), some protocols allow retest or require that exact threshold.

- **Pregnancy and Contraception:** Because many breast cancer patients are premenopausal, *pregnancy exclusion* is crucial. The patient must not be pregnant or breastfeeding. A serum or urine pregnancy test (see Lab criteria) is done at screening and throughout the trial. Additionally, patients of childbearing potential must agree to use effective contraception during and for some time after the trial. This criterion is typically explained in the consent and eligibility form. EMRs might not have a direct way to enforce this, but documentation is kept in research records. Often, a progress note will mention "Counseled on pregnancy avoidance – patient on oral contraceptive" or similar. This criterion is virtually universal for systemic treatment trials. Only women who are clearly not fertile (postmenopausal or surgically sterile) are exempt from the contraception requirement. The presence of this criterion in protocols doesn't vary – it's always there for applicable patients.
- **Performance and Comorbidity Data Representation:** In EMRs, many of these clinical criteria (PS, diagnoses) are structured as discussed (Conditions for comorbidities, Observations for ECOG, DiagnosticReports for echo etc.). However, checking them for trial screening is often a manual or semi-manual process: e.g., a research nurse will review the problem list for disqualifying diagnoses, check the latest cardiology report for EF, confirm the ECOG from last clinic visit, etc. Some advanced systems might use a rules engine to flag obvious mismatches (for instance, a decision support algorithm could warn "Patient's EF 45% < required 50%" if one tries to enroll). The underlying data is there: e.g. FHIR queries could retrieve the latest LVEF Observation and compare to threshold, retrieve all Conditions with "heart failure" or "MI" etc. In narrative documentation, one may see an eligibility checklist: "ECOG 0 ✔; No cardiac history ✔; LVEF 60% ✔; No active infection ✔; …" which effectively shows each criterion and yes/no.

## Diagnosis Criteria (Breast Cancer)

- **Pathological Diagnosis of Breast Cancer:** All trials require a **confirmed diagnosis of breast cancer**, typically invasive carcinoma of the breast. This must be proven by histopathology (usually a biopsy of the breast tumor or metastatic site). For early-stage trials it could be the surgical pathology from a lumpectomy/mastectomy; for metastatic, a biopsy of a lesion. The EMR contains this in pathology reports. A typical inclusion line is "Histologically confirmed adenocarcinoma of the breast." The Condition on the problem list might say "Infiltrating ductal carcinoma of breast, ER-positive, HER2-negative" (with codes for the neoplasm). Verification status would be confirmed (since pathology proof exists) [17]. The actual pathology **DiagnosticReport** details tumor type, grade, etc. Trials often require that the pathology report be available for central review or to confirm key markers (ER/PR/HER2). As structured data, the diagnosis can be coded with ICD (C50.* for breast malignancy) or SNOMED. In narrative, the oncologist note will have something like "Diagnosis: Stage IIIB invasive ductal carcinoma, Grade 3, ER 0%, PR 0%, HER2 3+." Those details bring us to subtype criteria.

- **Cancer Subtype (ER/PR/HER2 Status): Hormone receptor (estrogen and progesterone) status** and **HER2 status** are fundamental in breast cancer. Trials are usually designed for a particular subtype:
- **HER2-positive** breast cancer (overexpression or amplification of the HER2/neu gene) accounts for ~15–20% of breast cancers [27]. Many trials (targeted therapies, antibodies) are specifically for HER2+ patients. Thus, a common inclusion is "HER2-positive disease, defined as IHC 3+ or ISH amplified per ASCO/CAP guidelines." Conversely, trials for other therapies might explicitly require **HER2-negative** status (to exclude HER2+ patients who have different standard options). In EMR, HER2 testing results are stored either in pathology reports (often the initial biopsy report includes HER2 by immunohistochemistry or FISH). It may also be parsed into discrete data – some EHRs have a "HER2 status" field. FHIR could represent it as an Observation (e.g., LOINC 11055-9 "HER2 immunohistochemistry stain [Score]"). The value might be "3+" or "positive". There is also a binary qualitative often: positive or negative. The pathology **DiagnosticReport** will have a conclusion like "HER2: Positive (3+ by IHC) [27]." Trial screening involves confirming this result in the record. If a patient's HER2 status is unknown, they must be tested before enrolling in a HER2-driven trial.
- **Hormone Receptor (HR) status:** Similarly, trials often target **ER-positive (and usually PR-positive by correlation)** breast cancer (which is about 70% of cases) versus **triple-negative** (ER/PR-negative, HER2-negative) which is ~15%. A trial for a new endocrine therapy would require **ER-positive disease** (often defined as ≥1% cells positive for ER by IHC). A trial for triple-negative breast cancer (TNBC) requires ER<1%, PR<1%, and HER2-negative. These criteria come straight from pathology data. EMR pathology reports list ER % and PR %. Many EHRs also have discrete fields (yes/no for ER, PR). FHIR Observations could hold each (LOINC 41831-9 "Estrogen receptor [Presence] in Tissue by IHC", etc., or the percentage if quantitated). Checking eligibility means verifying the pathology report or summary: e.g. "ER 90%, PR 50%, HER2 0 – qualifies for ER+ HER2- trial" or "ER/PR 0, HER2 0 – qualifies for triple-negative trial." These biomarker criteria are among the most frequently encountered, since essentially every breast cancer trial will specify the subtype it's intended for.
- **BRCA mutation:** Some trials (especially PARP inhibitors or DNA damage response therapies) specifically enroll patients with **BRCA1 or BRCA2 germline mutations** or other mutations (PALB2, etc.). For example, trials for PARP inhibitors initially focused on BRCA-mutated metastatic breast cancer. This adds a genetic testing requirement for eligibility. Many breast cancer patients are tested for BRCA as part of routine care if young or TNBC, etc. If a trial needs it, the EMR might have genetic counselor notes or a genetic test report (usually stored as a PDF or in a genetics module). FHIR can represent a **FamilyMemberHistory** and genetic **Observation** for BRCA mutation (there are specific sequence variant IDs). The presence of a pathogenic BRCA1 mutation would be an inclusion criterion. This is not in a majority of trials, but in a subset focused on hereditary breast cancer (~it appears in those targeted trials and now some broader "BRCA or homologous recombination deficient" cohorts).
- **PD-L1 status in TNBC:** Some immunotherapy trials for triple-negative breast cancer require **PD-L1 positive tumors** (similar to lung). For instance, the IMPassion130 trial of atezolizumab in TNBC required PD-L1 expression on immune cells ≥1%. So certain trials now include PD-L1 as a criterion for TNBC patients. PD-L1 testing in breast cancer is not universal, but for those trials, a test (like the SP142 assay) is done. In EMR, that result would be a pathology addendum or separate report, recorded as an Observation (possibly similar to how it's done in lung, just different scoring). It might say "PD-L1 on immune cells: 10%, positive by trial criteria." This is a rising criterion in the immuno-oncology subset of breast cancer trials.
- **Disease Stage Requirements:** Trials clearly state whether they are for **early-stage** disease (non-metastatic, typically after surgery or neoadjuvant therapy) or for **metastatic** (Stage IV) disease. Inclusion criteria will specify something like:

- For adjuvant trials: “Histologically confirmed Stage II or III breast carcinoma that has been completely resected” or “no evidence of metastatic disease on baseline scans.” They may require certain tumor size or nodal involvement if targeting high-risk patients (e.g. “≥4 positive axillary lymph nodes” might be an inclusion for an aggressive adjuvant trial).
- For metastatic trials: “Unresectable locally advanced or metastatic breast cancer” and often “at least one site of metastasis that is measurable (if required)” and sometimes “if HER2+, must have progressed on standard HER2 therapy” (which crosses into prior treatment domain).

Stage information is derived from a combination of pathology and imaging. Early stage trials lean on pathology (surgical pathology report for tumor size, nodes) plus imaging to confirm no spread. Metastatic trials rely on imaging showing metastases. EMRs often have staging captured at diagnosis (AJCC stage recorded as text or in cancer registry). A FHIR **Observation** or **Condition.stage** field can hold the stage classification. For example, a Condition for breast cancer might have stage IV as a code. Many times, though, stage is noted in the doctor’s note (“Stage IV at diagnosis with bone and liver mets”). For eligibility, the coordinator checks those notes and baseline scan reports to ensure the patient fits the stage criteria. - **Disease-Free Interval (for early stage):** If a trial is enrolling patients after initial treatment (e.g., a trial of extended adjuvant therapy), they might require that the patient is currently without evidence of disease but at high risk for recurrence. Or if enrolling post-neoadjuvant, maybe they require residual disease pathologically. These specifics come from surgical pathology or recent scans. For example, “patients must have residual invasive disease >1 cm after neoadjuvant chemo” – that requires reading the pathology report measurement. If the EMR has synoptic reports, that could be a discrete element. Usually, it’s manually read by the investigator. - **Representation:** The diagnosis and staging criteria are captured in multiple structured ways: **Condition** resources for the diagnosis (with potentially stage and histology as attributes), and **DiagnosticReport/Observation** for biomarkers that define sub-diagnoses (ER/PR/HER2 status). Pathology reports and lab results carry those critical pieces [18] [19]. In unstructured form, oncology summaries often succinctly state these: e.g., “Triple-negative breast cancer, initially Stage III, now metastatic to lungs and liver.” That one sentence in a note covers subtype and stage. Trial eligibility reviewers use such summaries heavily. - **Central Confirmation:** Some trials (especially global ones) require central re-testing of HER2 or ER to confirm eligibility. This isn’t in the EMR initially – it happens after initial screening. But it means the site must send tumor blocks and a central lab will verify, for example, HER2 status. If a central test overturns the local result, the patient could be deemed ineligible. This underscores how crucial and sometimes finicky these pathology criteria are. Locally, though, the site uses the EMR’s pathology data to decide whether to even proceed with central testing.

## Treatment History Criteria (Breast Cancer)

- **Prior Therapies in Early-Stage Setting:** For **adjuvant (early-stage) trials**, eligibility often hinges on what prior therapy the patient has had:
- If it’s a trial of an adjuvant treatment vs placebo, typically the patient *must have completed standard surgery and (if indicated) chemotherapy/radiotherapy* before trial enrollment. Criteria may say “patients must have completed definitive surgery with clear margins and (neo)adjuvant chemotherapy if given, at least 4 weeks prior to enrollment.” Also “any toxicity from prior therapy must be ≤ Grade 1.” So essentially, the prior standard treatments should be done and the patient recovered. The trial might allow patients who got chemo vs those who didn’t depending on design, but they often stratify or require it for high-risk populations.
- If testing something like extending endocrine therapy, a trial might require that the patient has been on endocrine therapy for X years already. For example, “patients who have completed 5 years of tamoxifen are eligible to be randomized to extended therapy.” That’s prior treatment by definition.

- Some trials might exclude patients who *did not* get standard chemo if it was indicated (to avoid including low-risk patients). So they might say "must have received an anthracycline and taxane regimen if lymph nodes ≥4 positive" – basically ensuring all had similar baseline treatment.

These criteria rely on treatment records: surgery reports (Procedure in EMR), chemotherapy records (MedicationAdministration or Oncology flow sheet), radiation completion (Procedure). Often, the oncologist's note or discharge summary from initial treatment will list all that. The coordinator verifies dates (e.g. surgery date, chemo start/end dates). Structured data exists: many centers record chemo in a structured regimen module, and surgeries are documented as coded procedures. But assembling that sequence is complex; typically a manual timeline is made to ensure the patient meets any timing windows. - **Prior Lines in Metastatic Setting:** For **metastatic breast cancer trials**, specifying prior lines of treatment is extremely common: - A **first-line metastatic trial** (e.g., first therapy for metastatic disease) will often require "no prior systemic therapy for metastatic disease" (though prior adjuvant therapy is usually allowed). If a patient relapsed 2 years after adjuvant chemo, they still count as "no prior metastatic chemo," so they qualify. - A **second-line trial** might say "must have progressed after exactly one prior chemotherapy regimen for metastatic disease." Or if it's a hormone therapy trial, they might require "progression after at least one and no more than two prior endocrine therapies for metastatic disease." - Some trials that test something in heavily pretreated populations will require "at least 2 prior chemotherapy regimens in metastatic setting," etc. - There are trials for patients who have "triple-negative breast cancer with at least two prior treatments" etc.

Essentially, they're segmenting the patient population by how much prior treatment they've had. This is verified by reviewing the patient's medication history in the metastatic setting. EHR's medication list or oncology treatment history is directly relevant. In some cases, the line between adjuvant and metastatic therapy is blurred (e.g., if a patient relapses shortly after finishing adjuvant therapy, some trials count that therapy as a line in metastatic). The eligibility will define it clearly. In the EMR, one might find a chronological list: "Adjuvant chemo: AC-T completed 2018; metastatic chemo 1: capecitabine 2020; metastatic chemo 2: vinorelbine 2021;" etc. Many oncology systems have a summary of systemic therapies given, which could be output as structured data or at least as a table in notes. If FHIR resources were used, each regimen could be a **CarePlan** or each drug a **MedicationAdministration** with dates. To check eligibility, the coordinator counts those. If the criterion is "≤2 prior lines in metastatic setting," a patient with 3 lines would be excluded. - **Specific Prior Drugs Exclusion:** Trials often specify that patients **must not have received the same class of agent** that is being studied: - For example, a trial of a new CDK4/6 inhibitor in HR+ breast cancer might exclude patients who have had any CDK4/6 inhibitor previously (palbociclib, ribociclib, etc.). So "no prior CDK4/6 therapy" would be a criterion. - Trials of a novel HER2 drug may exclude anyone who had certain other experimental HER2 therapies (to have a cleaner population). - If a trial is combining two drugs, sometimes prior exposure to either is disqualifying.

These require detailed medication history checking. The EMR medication list (with generic names and dates) is crucial. For example, if a patient was in a prior trial of a similar drug, that might appear as "investigational drug X" in their record or not at all (if it was blinded). Often, the oncologist's note would mention it. So some detective work may be needed. But key meds like "palbociclib" would likely be in the med list or pharmacy records. In structured terms, one could search **MedicationStatement** for any with a certain ingredient code. In text, a note might clearly say "Progressed on palbociclib last year" – an immediate red flag for the new trial. - **Radiation and Surgery History:** Trials may have criteria about **previous radiation** or **surgery**: - If the trial therapy is localized (like a device or intra-operative treatment), prior surgeries might exclude or require a certain type. Generally for systemic trials, it's more about timing – e.g. "if patient had major surgery, must be >4 weeks recovered" (similar to lung criteria). - For radiation, if the trial is testing something in a previously radiated area or combining with radiation, etc., they might require "no prior radiation to the only measurable lesion" (so that lesion's

response can be assessed without confounding). - Some trials specifically exclude prior radiation to certain areas if expecting to use those as endpoints. Or if a trial drug interacts with radiotherapy, prior extensive radiation might exclude. - In adjuvant setting, typically “must have completed any necessary radiation therapy” is a criterion to avoid overlapping toxicities when starting trial treatment.

The EMR tracks radiation therapy as **Procedure** entries or dictated summaries (“completed whole-breast irradiation on 1/1/2022”). Checking these is straightforward from oncology notes or radiation oncology summaries. - **Time From Prior Therapy:** Similar to lung, **washout periods** are used: e.g. “≥2 weeks since last chemotherapy dose or radiation; ≥3 weeks since major surgery.” This ensures recovery. These are checked by looking at the dates of last treatment in EMR. If last chemo was given 10 days ago and criterion is 21 days, the patient must wait to enroll. EMR chemo administration records have exact dates; the coordinator calculates the interval. - **Outcome of Prior Therapy:** Some criteria require that the patient had a certain outcome to prior therapy. For example, a trial might require that a patient’s disease **relapsed during or after adjuvant therapy** (to ensure they have aggressive disease). Or that they **demonstrated progression on the most recent regimen** (to ensure they need a new treatment). Usually, any patient going into a trial has progressing disease by definition, but occasionally if a trial is maintenance or something, they might require stable disease. These are more nuanced and involve interpreting clinical notes and scan reports (not purely structured – though *progression* vs *response* might be captured in an tumor assessment Observation if RECIST evaluations are recorded). - **Representation:** Prior treatment data is partly structured (lists of meds, procedure codes for surgeries, radiotherapy records) and partly narrative (oncologist summary). On a FHIR server, one could query for **MedicationAdministration** of known chemo drugs, but one needs to know which ones to look for. Often, a simpler approach is looking at **Encounter** histories or **Procedure** logs for chemotherapy sessions. EHRs often have a chemotherapy module separate from the main medication list, which might not directly translate to FHIR without specific mapping. In any case, trial eligibility checks rely heavily on an oncology summary of what treatments the patient has had. Many oncologists include a “Treatment History” section in notes (which is unstructured but very useful). For example:

> Treatments: AC x4 (completed 2015), paclitaxel x12 (completed 2015), letrozole from 2016–2018 (stopped due to recurrence), capecitabine 2019 (progression), currently on vinorelbine.

From this one paragraph, one can deduce lines of therapy and exposures. That would be used to see if the patient fits the trial that, say, requires ≤2 prior chemotherapies for metastatic disease. - **Exclusion of Inadequate Response:** Some trials might exclude patients who had a very poor response to all prior therapy (implying very refractory disease), but typically if disease is too refractory, the trial might worry about futility. However, this is rarely an explicit criterion, more of a consideration for whether the patient should go on a trial or not ethically. There are also criteria like “no prior hypersensitivity reaction to similar drugs,” which come under treatment history (e.g., if a patient had a severe reaction to paclitaxel, they might exclude them from a trial of a similar taxane). That would be documented in allergy/adverse reaction lists in EMR (FHIR **AllergyIntolerance** resource could capture “paclitaxel anaphylaxis”).

## Imaging Criteria (Breast Cancer)

- **Brain Metastases:** Breast cancer, especially HER2+ and triple-negative subtypes, can metastasize to the brain. Historically, patients with active **brain metastases** were often excluded from general systemic therapy trials, similar to lung cancer. They would only be eligible after local treatment (surgery or radiation) and if stable neurologically. The Fred Hutch/ASCO study indicated a shift to be more inclusive: about 43% of trials allowed previously treated brain mets, while ~14% still strictly excluded any brain mets [6]. In breast cancer trials, this criterion is explicitly stated: e.g. “CNS metastases must be stable (no progression for ≥4 weeks) and patient

off steroids, or else patient is excluded." Many metastatic trials now allow treated/stable brain mets because therapies (like HER2 drugs) can have CNS activity. But if a patient has untreated or progressing brain mets, they'll be excluded due to poor prognosis and potential drug penetration issues. EMR detection of brain mets is via imaging (MRI brain reports). These would show up as Conditions (metastatic disease to brain) in the oncology problem list if the doctor documented it, or in radiology **DiagnosticReport** (impression: "metastatic lesions in brain"). To verify stability, sequential MRI reports are compared. This doesn't map easily to simple structured data aside from date comparisons of report text. It's assessed by the investigator reading radiology reports and neurologic exam notes. In FHIR, imaging can be represented, but whether a metastasis is present is an interpretative result, often just text. Some systems might code metastasis as a separate Condition with an onset date.

- **Measurable Disease:** For metastatic breast cancer trials that aim to measure tumor response, **measurable disease by RECIST 1.1** is usually required (except in certain trials focusing on progression-free survival regardless of measurability). Measurable disease in breast cancer often means at least one lesion in viscera or nodes ≥1 cm by CT, or a skin lesion that can be measured with calipers. Patients with only bone metastases (which often are unmeasurable lytic/blastic lesions on bone scan) can be a special case – some trials allow "bone-only disease" even if not measurable by RECIST, but others exclude them to not complicate response assessment. If a trial requires measurable disease, a patient with bone-only metastases would be ineligible. This criterion is present in a large fraction of metastatic trials (most phase II/III). It's handled the same way as in lung: by reviewing imaging findings. The EMR radiology reports and possibly oncologist's assessment note ("Patient has measurable liver lesions of 2 cm per CT"). There may not be an automated detection of "measurable disease present" – radiologists don't label something as measurable vs not in their standard reports. However, at many research sites, a baseline RECIST assessment is done: target lesions are identified and measured. These can be entered into a specialized system or sometimes an EHR module. Those target lesion data (sizes) could be stored as Observations. But generally, initial eligibility regarding measurability is determined by a radiologist or clinician reading the latest scans. For structure, one could say if any tumor dimension in a **TumorMeasurement** observation is ≥10 mm, then measurable disease exists. Such Observations might be recorded in a cancer registry or trial management system rather than the EHR proper.
- **Baseline Imaging for Staging:** In early-stage breast trials (adjuvant), there's usually a requirement that a baseline workup showed no metastatic disease. For instance, "no evidence of metastatic disease on chest/abdomen imaging and bone scan." This is implied by the trial's focus on early disease. So before enrolling, patients often get scans to confirm cancer hasn't spread. If a scan unexpectedly found metastasis, the patient would switch to a metastatic trial instead (and be ineligible for the adjuvant trial). Thus, imaging criteria in early disease ensure the correct staging. It's not always written as a criterion ("exclusion: metastatic disease present"), but it's inherent. Checking this uses the diagnostic imaging reports in EMR (e.g. a PET/CT report – if it's clean, proceed; if not, exclude).
- **Ejection Fraction Imaging: Cardiac imaging** is effectively an eligibility test in HER2 trials. We covered LVEF under clinical, but the actual measurement comes from imaging (echocardiogram or MUGA scan). So one could view it here: The criterion "LVEF ≥50%" means the patient must undergo an echo/MUGA and the imaging **DiagnosticReport** must show EF high enough. If an EF is borderline, sometimes repeating the imaging is allowed once to see if it improves (accounting for variability). In FHIR, this is an Observation (quantitative). In text, the investigator might write "Baseline MUGA LVEF = 48%, ineligible due to low EF" if that happens.
- **Other Organ Imaging:** If a trial has unique imaging needs (for example, a trial of a bone-targeted radiotherapy might require at least one bone metastasis visible on bone scan), that becomes a criterion. Or a trial of liver-directed therapy might require a measurable liver lesion

specifically. These are specialized cases. The EMR's imaging results would again be consulted for those specifics (e.g. "bone scan shows multiple bone mets" or "CT shows liver lesion").

- **No Extensive Prior Radiation to Measurable Lesions:** Some protocols say if the only measurable tumor was previously irradiated, it cannot be used as a target unless progression is documented there. So effectively, if all measurable disease has been treated by radiation (and hence is either scar or stable), the patient might be excluded for lack of evaluable disease. This is a nuance often handled in trial site meetings.
- **Central Imaging Review:** Occasionally, eligibility might hinge on a central radiology review (for example, to confirm the presence of a specific lesion type or brain met status). That means local EMR info is preliminary, final eligibility is confirmed by central read. But that is part of trial logistics rather than EMR content.
- **Representation:** Imaging results are stored as narrative reports in most systems, though key findings can be codified. FHIR **DiagnosticReport** can carry the text and also reference to images (DICOM) if needed [29]. Some advanced use-cases create **Observation** entries for each significant imaging finding (like an AI algorithm might label metastases), but typically not done systematically. So trial eligibility for imaging factors remains a manual process of reading radiologist impressions. There is some move toward **synoptic radiology** reporting which could structure things like "Brain mets: present/absent" but that's not widespread. For now, the investigator's interpretation is the gold standard documented in a clinic note: e.g. "CT 6/1/2025: liver and lung mets present (measurable), no brain mets on MRI – patient meets imaging criteria."

## Pathology/Biomarker Criteria (Breast Cancer)

- **ER/PR and HER2 Status (Biomarker Recap):** As mentioned, **ER, PR, HER2** are crucial biomarkers essentially defining separate diseases within breast cancer. Trials nearly always specify one of the categories:
- Trials for **hormone receptor positive, HER2-negative** breast cancer (this is the largest subset) – e.g., endocrine therapy trials, CDK4/6 inhibitor trials, etc. Inclusion: "ER-positive (≥1%) and/or PR-positive, and HER2-negative breast cancer."
- Trials for **HER2-positive** breast cancer – all HER2-targeted therapy trials. Inclusion: "HER2-overexpressing tumor (IHC 3+ or FISH-amplified)." They usually also require ER/PR status, either not important or stratified (some HER2 trials also require HR-negative specifically or allow both).
- Trials for **Triple-Negative** breast cancer – often immunotherapy or chemo trials. Inclusion: "ER <1%, PR <1%, and HER2 0 or 1+ (if 2+, FISH not amplified)."

These criteria are the top-occurring terminologies because essentially every trial will mention "HER2-positive" or "HER2-negative," etc. As open-source evidence, one can note that ~15-20% of breast cancers are HER2-positive [27], so a significant fraction of trials (and nearly all HER2 drug trials) are devoted to that group, making "HER2-positive" a frequent eligibility term. Similarly, ER-positive defines the majority group. In EMR, these markers are usually determined at initial diagnosis, but might be retested on metastasis (especially HER2 or ER can sometimes change). The **Pathology DiagnosticReport** and subsequent **Observation** entries capture these. For example, a pathology report might include:

> *Immunohistochemistry:* ER: 95% nuclei positive, Allred score 8; PR: 10% positive, Allred 4; HER2: 3+ (Positive).

This single report provides all needed info for eligibility categorization. The trial site will often transcribe this into their eligibility checklist. If multiple biopsies were done, the most recent or metastatic site's results might be used per protocol (some trials require confirmation that metastatic lesion is also HER2+, etc., if there's discrepancy). - **Molecular Subtypes and Genomic Markers:** Beyond the classical markers, new trials look at genomic tests: - **BRCA1/2 mutations:** As mentioned, certain trials (PARP

inhibitors like olaparib, talazoparib) initially targeted germline BRCA-mutated patients. So "Known deleterious or suspected deleterious BRCA1 or BRCA2 germline mutation" was a key inclusion criterion in those studies (e.g., OlympiAD trial for olaparib in BRCA-mutated breast cancer). Now, some trials include not just BRCA but any **homologous recombination repair (HRR) gene mutation**. This broadening means they might include PALB2, CHEK2, etc. Still, BRCA is the main one. If the patient hasn't had genetic testing, the trial might require it to determine eligibility. The EMR might contain a genetic testing **DiagnosticReport** with results: e.g., "BRCA1 c.68_69delAG pathogenic mutation identified" – which in FHIR would be an Observation with that variant info. If a patient is known BRCA-positive, it might also be in the problem list ("BRCA1 mutation carrier" as a Condition). - **PIK3CA mutation:** In hormone receptor positive breast cancer, trials for PI3K inhibitors (like alpelisib) required tumor **PIK3CA mutations**. The SOLAR-1 trial, for example, only enrolled those with a PIK3CA mutation. So that's another biomarker criterion that appears – often determined via tumor sequencing (PCR or NGS panel). The pathology/molecular lab report would list PIK3CA gene status. Many centers run gene panels on metastatic tumors, so the EMR might have an NGS report listing dozens of genes with PIK3CA highlighted as mutant or not. In FHIR, that can be an Observation with an identified variant. If a patient's report shows PIK3CA E545K mutation, they meet that criterion. - **PD-L1 (in TNBC)** – discussed earlier; it's an IHC biomarker. In the context of, say, atezolizumab plus chemo in TNBC, eligibility required PD-L1 immune cell stain ≥1%. That's gleaned from a pathology report addendum. Not all TNBC patients are tested for PD-L1 outside trials, but for trials it's done. If tested in the hospital lab, the result is stored similarly to other IHC results. - **Genomic assays (prognostic):** In adjuvant settings, tests like Oncotype DX (21-gene recurrence score) categorize risk. Some trials for adjuvant therapy have used these to select patients (e.g., the TAILORx trial used Oncotype to decide randomization). A criterion could be "Oncotype recurrence score 11–25" for example. That means the patient had that test done. Those assay results often come as PDFs attached to EMR. If integrated, they might be a PDF in a media resource. Rarely structured because these commercial assays are not always interfaced. But they could be manually entered as an Observation (with code for "Recurrence Score" and the numeric value). The trial eligibility is directly numeric from that report. This is a specialized case, but notable because it was used in trials to identify who might not need chemo, etc. - **Pathology Grade and Features:** Occasionally a trial might specify tumor grade or other histologic features (e.g., trials specifically for "triple-negative, androgen-receptor positive" breast cancer have been explored – requiring AR IHC positive in TNBC). Or a trial for a specific rare subtype, like "metaplastic breast carcinoma only," which is a histologic variant. Those criteria are less frequent but when present, rely on pathology report text (which in structured form could be coded as histologic subtype in SNOMED). For example, SNOMED codes exist for metaplastic carcinoma; if one had those coded, you could query for patients with that code. - **Pathology Node Status (early stage):** Some adjuvant trials pick patients based on how many lymph nodes were positive. E.g. trials for node-positive disease might say "≥4 positive axillary lymph nodes" or conversely, some trials for lower risk might require 1-3 nodes positive. That information is in the surgical pathology report (e.g. "10 out of 15 lymph nodes involved"). That might not be structured except as a narrative or synoptic section. Some pathology systems provide discrete counts which could be pulled as Observations. But mostly it's read by humans. If required, the coordinator sees "N2 disease (4 nodes)" and knows it fits a criterion "≥4 nodes." FHIR could model number of positive nodes as an Observation (some standard exists for TNM staging values). - **EMR Representation:** As with lung, pathology and biomarker results in breast cancer are captured via **DiagnosticReport** (pathology) and associated Observations for things like receptor status [18] [19]. Many institutions maintain a synoptic report for ER/PR/HER2 which can be parsed. For instance, the CAP eCC (electronic Cancer Checklists) output can be in discrete fields. In FHIR, one might see a BreastCancerProfile Condition with an extension for ER status. The CIMI/HL7 oncology models have detailed ways to record receptor status as part of the cancer diagnosis. But without going into standards detail, suffice it to say the data is there in some form and is critical for trial matching. Clinicians rely on pathology reports (unstructured or semi-structured) to confirm eligibility. It's routine for a research nurse to have a copy of the path report in hand when determining if a patient qualifies for a subtype-specific trial.

## Laboratory Criteria (Breast Cancer)

Breast cancer trials use lab eligibility criteria very similar to lung cancer trials (and oncology trials in general), adjusting for any specific drug toxicities. The typical lab cut-offs for **hematologic, liver, and renal function** are essentially identical to those described for lung. The table above shows frequencies: e.g., 84% of breast trials had bone marrow criteria, 51% had strict liver limits, 42% strict renal [3]. We will highlight any differences or special cases for breast cancer:

- **Hematologic:** The exact thresholds (ANC ≥1500/μL, platelets ≥100k, Hb ≥9–10 g/dL) are used across cancer trials and breast is no exception. One minor difference: if a trial involves a myelosuppressive therapy, they may be more stringent (e.g., require Hb ≥ 10 if expecting anemia, or maybe allow certain baseline anemia if they plan to transfuse). But generally the same. For breast cancer patients who have had heavy prior chemo, counts can be an issue (e.g., bone marrow reserve might be lower after multiple chemo lines). Trials sometimes slightly relax criteria for heavily pretreated populations (like allow ANC ≥ 1.0 if the drug is not too myelosuppressive). This is protocol-specific. The EMR lab data representation is the same as described: a CBC panel Observation. The **DiagnosticReport** for hematology might show everything, and Observations for each result with reference ranges. Checking is done by looking at recent labs (often within 1 week of enrollment).
- **Liver Function:** Similar cut-offs (AST/ALT, bilirubin) apply. One thing in breast cancer metastatic patients: if they have liver metastases, they might have moderately elevated AST/ALT. Trials generally allow up to 5× ULN if liver metastases are present (which is consistent with lung). So a patient with metastases causing AST 4× ULN is still eligible by that rule, whereas someone without liver metastasis wouldn't be. This distinction is often explicitly written. The EMR lab logic to apply here is: if Condition = liver metastases (yes), then threshold = 5× ULN, else 2.5×. That's not automated; it's handled by physician judgement. They see the patient's scans (liver mets Y/N) and labs, then interpret criteria accordingly. Some advanced decision support could incorporate it by linking the presence of a "liver metastases" diagnosis to criterion logic, but that's rarely implemented.
- **Renal Function:** The usual creatinine or CrCl ≥60 mL/min (some protocols say ≥50) is applied. Breast cancer patients, especially older ones, might have borderline creatinine clearance – this can exclude some older women. Like lung, this criterion eliminates a chunk of patients (in one general oncology study, strict CrCl cut-off of 60 would preclude 20–46% of real-world patients [30]). So it's a known restrictive factor. EMR labs for creatinine and eGFR we described. Often, an oncologist will calculate CrCl for an older patient to see if they can make it; if it's close, sometimes a hydration or retest could slightly improve it.
- **Cardiac Markers:** As discussed under clinical, for HER2 therapy trials, **LVEF measurement** is almost treated like a "lab" in that it's a quantitative criterion. It's actually imaging, but often grouped with baseline tests. For example, the screening checklist might list LVEF next to labs. We've covered representation of LVEF (echo reports).
- **Endocrine Labs:** Not typically required, but if a patient has, say, hypothyroidism, as long as it's controlled on meds it's fine. If it's uncontrolled (TSH extremely high or low), the general "no uncontrolled illness" could apply. But there's no standard trial criterion about thyroid for breast trials (unlike some immunotherapy in other diseases that require TSH monitoring).
- **Glucose and Lipids:** Some newer targeted therapies (like AKT inhibitors) can cause hyperglycemia or high lipids. If a trial drug has that known toxicity, they might require baseline **HbA1c below a certain level** or **cholesterol/triglycerides below certain levels**. For example, trials of AKT inhibitors (which can raise blood sugar) often required **HbA1c < 8%** and well-controlled diabetes to enroll. That's a specific criterion for those trials. Similarly, if a drug affects lipids, they might require cholesterol < some value or patient on treatment for it. These are

specialized but worth noting. EMR has HbA1c as a lab Observation, and lipids as well. It's unusual in oncology to have those criteria, but not unheard of for metabolic side-effect-prone drugs.

- **Bone-specific Labs:** If a trial drug has bone effects (like some bone-targeted radiopharmaceutical), they might require calcium levels normal, etc. Generally, it's covered by comprehensive metabolic panel anyway. Calcium or other electrolytes should not be dangerously abnormal. If a patient had, say, hypercalcemia from bone mets, that typically would exclude them until corrected (and an investigator would treat that as part of clinical care first).
- **Pregnancy Test:** Absolutely required in any trial with women of childbearing potential, as previously noted. Breast cancer trials often involve women up to middle age, so pregnancy is possible for some. The testing is done at baseline and periodically. It's documented in labs as an Observation (HCG test). This is identical to lung context. For breast cancer, sometimes even postmenopausal women get tested if the protocol doesn't trust patient report (some protocols define women > a certain age or without menses for 1 year as not needing test, others test everyone under 55, etc.). That's an operational detail.
- **Lab Data Usage:** The high prevalence of lab criteria means that often a screening lab panel is ordered for a trial candidate. Those results go in EMR and are scrutinized. If anything is out of range, the team may intervene (e.g., give a transfusion for low Hb, recheck labs after an infection resolves, etc.) to make the patient eligible. This dynamic isn't captured in static data – but it's an important part of eligibility: some criteria are "fixable" (like lab abnormalities) given time or treatment, whereas others (age, prior therapies) are not. That's why labs may be re-measured to get a patient in range if possible.
- **Data Format:** To reiterate, labs are structured and standardized, making them easily comparable to criteria. They are typically in SI or conventional units with reference ranges. Trials define cut-offs in terms of ULN (upper limit of normal). The EMR knows the ULN from the lab's reference range. So a smart system could automatically say "bilirubin 2.0 mg/dL, local ULN 1.2 mg/dL, that's 1.67× ULN, criterion is ≤1.5× ULN, so fails." This would require reading the Observation.value and Observation.referenceRange from FHIR and doing the math. Usually humans do this mentally or with simple calculations.
- **Frequency in Protocols:** The BMC Medicine study [3] shows just how common these lab-based exclusions are in breast trials (very high percentages). That aligns with anecdotal experience: nearly every oncology trial has a page of lab criteria. They may vary slightly in thresholds but not much.
- **Special Lab Criteria Example:** If a trial involves a drug that affects the liver, sometimes they include a criterion about **no active hepatitis** beyond just serology – e.g., requiring normal baseline viral loads for hepatitis B/C carriers or requiring patients with Hep B to be on antivirals. These are more detailed. For example, a trial might allow a Hep B surface antigen positive patient if they have DNA undetectable on antivirals (some trials say that now). That introduces lab monitoring of viral DNA. It's beyond the basic criteria, but shows how criteria can evolve to be more inclusive (instead of outright excluding all hep B carriers, they include them if labs show controlled infection). EMR lab data for HBV DNA, etc., would be Observations too (with specific LOINC codes).
- **Conclusion on Labs:** They ensure patient safety and homogeneity and thus remain a critical structured component of eligibility. Breast cancer trials mirror lung cancer in applying these across the board, with the addition that breast-specific treatments (HER2 drugs) bring in a requirement for imaging-derived "lab" like LVEF.

## Alzheimer's Disease Trials

**Overview:** Clinical trials in Alzheimer's disease (AD) and other dementias have very different eligibility focus compared to oncology. AD trials emphasize demographics (older age ranges), cognitive and functional **clinical assessments** (e.g. MMSE scores, stage of dementia), **neurologic and psychiatric**

**comorbidities** (excluding other brain disorders), and often require evidence of Alzheimer's pathology via biomarkers (amyloid PET or cerebrospinal fluid). Inclusion criteria typically ensure participants have a diagnosis of probable Alzheimer's at a defined stage (e.g. mild cognitive impairment or mild dementia), with generally good physical health aside from their cognitive impairment. Exclusions often rule out alternative causes of dementia and conditions that could interfere with study procedures (like MRI scans or lumbar punctures). **Table 3** gives common criteria in AD trials and how often they appear. We then describe details by domain and EMR data considerations.

**Table 3. Common Alzheimer's Disease Trial Eligibility Criteria (Frequency in Protocols)**

| Domain | Common Eligibility Criteria – Alzheimer's Disease | Prevalence in Trials |
|---|---|---|
| Demographics | **Age range** (e.g. 50–85 years for AD dementia trials) | ~87% of AD trials use age cutoffs [31] (typically mid-adulthood to elderly) |
| | **Language**/education: e.g. must speak English (for cognitive testing) | ~42% require English fluency [31] (in US-based trials) |
| | Availability of a **caregiver/study partner** (able to attend visits) | Very common (essential in most AD trials [32], as patient+partner "dyad" needed) |
| Clinical Status | **Cognitive status**: impaired but within a specific range (e.g. Mini-Mental State Exam **MMSE 20–26** for mild AD) | Nearly all trials – specific score ranges (varies by stage: mild, moderate, etc.) [33] |
| | **Diagnosis of Alzheimer's** (per clinical criteria, NIA-AA or DSM) | Universal (all trials confirm probable AD diagnosis) [34] |
| | **Functional status**: e.g. able to perform basic ADLs with assistance | Common inclusion (ensures not too severe; often implicit in staging criteria) |
| | **General health**: stable medical condition aside from dementia | Common – exclude uncontrolled chronic illnesses or frailty (most trials) [31] (65% exclude specific neurologic, 61% psychiatric disorders) |
| Diagnosis | **Probable AD dementia or MCI due to AD** (per standardized criteria) | Universal (explicit diagnostic criteria required) |
| | **Exclusion of other neurodegenerative diseases** (e.g. no Parkinson's, stroke, FTD) | Very common (65% trials specify neurologic exclusions [31]) |
| | **Disease stage**: e.g. mild cognitive impairment (CDR 0.5) or mild dementia (CDR 1) as required by trial | Specified in all trials (target population defined by stage) |
| Prior Treatment | **Cholinesterase inhibitor (AChEI) use stable** (if allowed) or no AChEI | Common: many trials require stable dose of donepezil or similar for ≥3 months, or disallow if not on one (depends on study design) |

| Domain | Common Eligibility Criteria – Alzheimer's Disease | Prevalence in Trials |
|---|---|---|
| | No investigational AD drugs or recent trial participation | Common in most trials (no other trial within e.g. 60–90 days) |
| | No psychoactive meds above certain doses (antipsychotics, benzodiazepines) | Many trials exclude heavy sedative use (to avoid confounding cognition) |
| Imaging | **Brain MRI** with changes consistent with AD and no significant other pathology | Nearly all trials require an MRI to exclude strokes, tumors, etc. (exclusion of major vascular lesions, etc.) |
| | No **extensive cerebrovascular disease** on MRI (e.g. excessive white matter lesions, multiple strokes) | Very common exclusion (part of "no other neurologic disease" check via imaging) |
| | No **large brain abnormalities** (e.g. tumor, normal-pressure hydrocephalus) | Common exclusion (via MRI screening) |
| | Tolerance of MRI/PET: no severe claustrophobia, MRI-incompatible implants | Common (explicitly exclude those who cannot undergo imaging) [35] |
| Pathology/ Biomarker | **Amyloid biomarker positive** (amyloid PET scan or CSF Abeta42/ tau) | Increasingly common (modern disease-modifying trials often require; e.g. majority of Phase III DMT trials) [36] |
| | **APOE4 genotype** – not usually required, but may stratify dosing (some trials obtain genotype for safety) | Some trials involve APOE genotyping (not as a yes/no criterion generally, except in targeted prevention trials) |
| | **Exclusion of other dementia biomarkers** (e.g. normal pressure hydrocephalus signs, extensive vascular changes) | Common (part of excluding other diagnoses, often via imaging/CSF) |
| Laboratory | **Rule out reversible causes**: e.g. B12 level, thyroid (TSH) normal | Very common – required workup: B12, TSH must be within normal or treated (standard in diagnostic criteria) |
| | **General labs**: acceptable CBC, liver, renal function (to ensure overall health) | Common (baseline safety labs, but thresholds often more lenient than oncology) |
| | **Negative for other brain-affecting conditions**: e.g. negative syphilis serology, HIV status (if cognitive impairment could stem from these) | Many trials require these tests to exclude other causes (not always explicitly stated, but often done in workup) |

| Domain | Common Eligibility Criteria – Alzheimer's Disease | Prevalence in Trials |
|---|---|---|
| | **No uncontrolled diabetes or hypertension** (as they affect cognition) | Common exclusion if severe (part of general health criteria) |
| | **Pregnancy test** (for women, though most are post-menopausal) | Required in trials if women of childbearing potential are included (many AD patients are older women, often post-menopause, but younger early-onset trials would test) |

## Demographic Criteria (Alzheimer's)

- **Age Range:** Alzheimer's trials typically set an **age range** that aligns with the population of interest. Most commonly, trials for late-onset AD recruit older adults, often **age 50 or 55 up to around 85** years [34]. For example, a trial for mild Alzheimer's dementia might include ages 55–85. Early-onset AD trials might include people as young as 40 if targeting genetic AD. The upper age limit (often mid-80s) is to exclude frail individuals who might have higher risk of complications or competing health issues. A study of 196 AD trials found that **87% had age cutoffs** in their eligibility criteria (often both a minimum and maximum age) [31]. In EMR, age is readily known from birthdate (structured in Patient resource). So checking age is straightforward: e.g., a patient aged 90 would be excluded from a trial capped at 85. The presence of age limits means trial participants are not extremely old on average, which has been noted as a generalizability issue (real-world AD patients often are >85). Still, most trials cap around mid-80s for safety and practical reasons (e.g., higher likelihood of frailty or dropout beyond that). Age criteria appear prominently in trial listings ("Inclusion: Age 50–85 years") and are non-negotiable.
- **Language and Education:** Many AD trials require that participants (and sometimes their caregiver) can **speak and read the language of the test instruments** (e.g., English in U.S. trials). This is because cognitive assessments like the MMSE or ADAS-Cog are validated in certain languages. In the cited analysis, 42% of trials explicitly restricted to English-speakers [37], which likely reflects U.S.-based studies. Additionally, some trials implicitly require a certain **educational level or literacy**, since testing memory and other cognition is hard if someone is illiterate. While not always a hard criterion, some protocols exclude those with very low education if the tests can't be administered validly (or they adjust normative data for education). In practice, education is captured in research forms but not always in the standard EMR (though sometimes recorded in social history). EMR might mention occupation or education level in a note. Language is typically noted by the clinical team ("primary language: English") as a patient attribute (could be in Patient.communication in FHIR). For eligibility, teams ensure the patient can complete evaluations in the required language; if not, unfortunately they're not included (which is a barrier to diversity).
- **Study Partner/Caregiver Availability:** A unique and **crucial inclusion criterion in AD trials is the requirement of a reliable caregiver or "study partner."** AD trials "require not one but two participants: the patient and a study partner" [32]. This person, often a family member, must spend significant time with the patient (e.g. "at least 10 hours per week with the participant" [35]) and be able to accompany them to study visits. The caregiver provides collateral information on the patient's daily functioning and helps ensure protocol compliance. Thus, almost all AD trials list something like "**presence of a responsible caregiver** (spouse, adult child, etc.) who can attend visits and monitor the patient's compliance" as an inclusion criterion. If a patient lives alone with no available caregiver, they will likely be ineligible [32]. In the EMR, caregiver

information might be found in clinic notes or a designated contact person field. It's not usually structured beyond emergency contacts. But during screening, sites will explicitly document caregiver name, relationship, and their agreement to participate. Some trials even have the caregiver sign consent (because they themselves might be asked to do tasks or fill out questionnaires). So effectively, "has a caregiver who can participate" is a yes/no criterion that's checked through patient interview. Research staff may note "Patient's wife will serve as study partner – meets criterion." If no suitable partner, that's a screen failure reason.

- **Consent Capacity:** While not always stated in criteria, it is implied that the patient must **be able to give informed consent** (or in some cases, if they're too impaired, a surrogate can consent depending on local laws and trial design). Many AD trials focus on mild stages where patients can consent. For moderate/severe AD trials, protocols often require consent from a legally authorized representative and assent from the patient. This is more of a regulatory requirement than a typical criterion bullet, but effectively, if a patient is too impaired to assent or has no surrogate, they cannot enroll. In EMR, cognitive status is noted; formal capacity assessments might be documented.
- **Sex:** Both men and women are affected by AD, and trials generally include both. There is no sex restriction (aside from needing to handle pregnancy risk if any – which is rare since participants are usually older). If a trial involves a drug with potential risk to a fetus (like some antibody with unknown risk), they might require women of childbearing potential to test negative for pregnancy (some early-onset AD trials could involve women in their 40s who could technically get pregnant). But the typical AD trial population is post-menopausal women and older men. So sex isn't a limiting factor. In EMR, sex is in demographics. We mention pregnancy tests under labs if applicable – usually for AD patients above 55 it might not be required, but if someone is e.g. 50 and female, they would likely still do a precautionary test.
- **Ethnicity and Genetic Background:** Not usually criteria, but some trials focusing on specific populations might have additional enrollment targets (like a study specifically in APOE4 carriers might indirectly specify that genotype, but that's biomarker, not demographic). There have been prevention trials targeting people with certain ancestry if they have higher risk of autosomal dominant AD (e.g. Colombian kindred), but those are specialized. In general, race/ethnicity are collected but not used to include/exclude except ensuring materials in language patient understands.
- **Representation:** Demographic info like age and sex and language lives in the EMR's **Patient** record (FHIR Patient.birthDate, gender, and possibly an extension or Patient.communication for language). Caregiver presence is not a standard EMR field, though sometimes listed as next of kin. More often it's in notes. Some EHRs have care team contacts (like listing a caregiver as a personal representative). The criterion essentially requires an interview to confirm. There's no automated check for "caregiver exists." The importance is huge: lack of a study partner is a major reason for screen failure in AD trials, contributing to the difficulty in recruitment [32] .

## Clinical Criteria (Alzheimer's)

- **Cognitive Status and Stage (MMSE, etc.):** One of the most defining inclusion criteria is the **cognitive test score range**. Trials choose participants at a certain stage of cognitive impairment:
- **Mild Cognitive Impairment (MCI)** trials: usually target individuals with memory impairment but not dementia. Criteria might include a global CDR (Clinical Dementia Rating) of 0.5, and an MMSE in the high 20s (perhaps 24–30) or MoCA (Montreal Cognitive Assessment) in a certain range. For instance, an MCI trial could require MMSE 24–30 inclusive (since MMSE max is 30, 24 is mild impairment).

- **Mild Alzheimer's dementia** trials: often require an MMSE roughly **20–26** (since ≤26 is often considered mild dementia) [38] . Some use MOCA or other scales similarly. They also may require a CDR of 0.5 or 1 (0.5 = very mild, 1 = mild).
- **Moderate AD** trials: might have MMSE roughly **14–24** or 15–20 depending on definition (moderate often 10–20).
- **Severe AD** trials: could allow MMSE <10 (some say 0–10 or 0–12).

Essentially, trials set cutoffs to target the right severity: not too mild (so they can detect decline) and not too severe (so they can still cooperate and have a chance to benefit). The SIU presentation snippet gives an example: mild studies MMSE 25–30, moderate 14–24, severe 0–13 [38] , though typically I've seen mild AD defined a bit lower (20–26). Regardless, every trial will specify a cognitive test score range. **Nearly 100% of AD trials use some cognitive inclusion criterion.** The Scientific Reports analysis didn't explicitly list this in the top 3 (since it's ubiquitous, not variable), but it's always there.

In the EMR, **cognitive test scores** might be found in neurology or neuropsychology notes. If a formal neuropsych battery was done, it's often scanned or summarized in a consult note (structured testing data seldom makes it to discrete EHR fields). However, something like MMSE might be done in clinic and recorded as an observation vital sign (some EHRs let clinicians record MMSE score as part of exam, or MOCA). In FHIR, a cognitive test result can be an **Observation** with a specific LOINC code (e.g., LOINC 72106-1 for MMSE score). If not structured, it's usually in the note ("MMSE = 23/30 on 7/1/2025"). Trial screening often involves administering an MMSE again to verify eligibility. Many protocols actually do the screening cognitive test as part of study screening, using their rater and standardized method, to ensure consistency. But they'll still often require that at baseline the score is in the range. So in terms of EMR data, the team might do an MMSE at the screening visit (which might be recorded on a source document and not necessarily in the EHR). Or if done, maybe in a research flowsheet or simply scanned.

Additionally, **global function scales** like CDR (Clinical Dementia Rating) or FAST (Functional Assessment Staging) might be used. Often inclusion says "Clinical Dementia Rating = 0.5 or 1" meaning very mild or mild dementia. CDR is based on caregiver interview. It might not be recorded in the EHR regularly, except if a neuropsychologist or memory clinic note includes it. Possibly as an Observation too, but typically just narrative "CDR = 1 (mild dementia)". The trial site will evaluate CDR at screening if required.

- **Diagnosis of Alzheimer's Disease:** Trials require that patients meet standardized criteria for **probable Alzheimer's disease** (or MCI due to AD). This usually means the patient has had a thorough clinical evaluation to rule out other causes of dementia, and their pattern of cognitive decline is consistent with AD. Many protocols cite NIA-AA (National Institute on Aging – Alzheimer's Association) criteria or DSM-5 criteria for Major Neurocognitive Disorder due to AD. Practically, this is confirmed by the study doctor. It's an inclusion that "the patient has a diagnosis of Alzheimer's disease" of a certain stage. Often documentation from the patient's physician or a workup is needed. In EMR, the **Condition** might be listed as "Alzheimer's disease" or "Dementia, Alzheimer type". This is the structured diagnosis, which ideally is present if they've been treated for it. If someone only has "dementia" unspecified in their chart, the trial may need to confirm AD with further tests (like amyloid PET).

So basically universal: being diagnosed with the condition under study is required, which is inherent but explicitly stated in AD trials. - **General Health and Medical Comorbidities:** AD trials often include a blanket statement that participants should be in **generally good health aside from their cognitive impairment**. They specifically want to exclude those with: - **Other neurologic diseases:** The criteria usually list **no other significant neurologic disease** that could affect cognition or safety. This includes **Parkinson's disease**, **stroke (CVA)**, **vascular dementia**, **Lewy body dementia (if not the target)**, **brain tumor, epilepsy** (if uncontrolled), etc. The analysis found 65% of trials explicitly listed some

neurologic exclusions [31]. For example: "Key exclusion: any history of stroke, significant head trauma, or other neurologic condition (e.g. MS, Parkinson's) that could cause cognitive impairment." If someone had a big stroke or has mixed dementia, they're out. In EMR, a history of stroke would be a Condition (ICD code for stroke, or imaging reports showing infarcts). Investigators will review MRI scans for silent infarcts; if more than a certain amount, that could exclude (addressed under Imaging). Similarly, if a patient has Parkinson's diagnosis in their problem list, that's an automatic exclusion because they might have Lewy body dementia or atypical parkinsonian syndrome causing cognitive issues. These diagnoses in structured form (Condition list: Parkinson disease code, etc.) and unstructured (neurologist notes) are used to determine eligibility. - **Psychiatric disorders:** Active and severe psychiatric illness is usually an exclusion because it can affect cognition and the ability to participate. The analysis reported 61% of trials had psychiatric exclusions [31]. Typically: **no uncontrolled major depression, schizophrenia, or bipolar disorder**. Mild depression and anxiety are common in AD and often allowed if stable on treatment, but if someone has, say, a history of schizophrenia (where cognitive deficits might be from that or medications), they'd be excluded. Also, **no current substance abuse** or alcohol abuse is a criterion (which can cause cognitive impairment). EMR data: psych conditions might be in problem list (Depression, etc.) and medication list (antidepressants, antipsychotics – which if high doses might exclude a patient or indicate a disorder). They also often require a depression screening scale (like GDS – Geriatric Depression Scale – below a certain cutoff) to ensure depression isn't severe. That might be done at screening as an observation. But baseline, they often need to have stable mood. In clinic notes, psychiatrists or PCP notes might mention if depression is present and treated. - **Systemic illnesses:** AD trials exclude those with **uncontrolled chronic diseases** that could interfere or increase risk. For example, **uncontrolled diabetes** (if someone's blood sugar is very high, they might have cognitive effects or risk in trial), **uncontrolled hypertension** (stroke risk), **severe cardiac disease** (like recent heart attack or heart failure that could shorten lifespan or preclude MRI if they have certain devices). They typically want someone who could likely complete the trial and not have another illness confound results or cause dropout. So inclusion often says "generally healthy for age, with no unstable medical conditions." Many protocols give specific exclusions: e.g. "no myocardial infarction in last 2 years, no unstable angina, no Class IV heart failure, no severe pulmonary disease requiring oxygen," etc. They might also exclude cancer if life expectancy is limited (<2-3 years). Since AD trials often run 18 months or more, they want people likely to survive that long. In EMR, these conditions are known via problem lists and doctor notes. It's somewhat subjective what counts as "well-controlled" vs not; investigators decide. For structured data, certain labs (like HbA1c for diabetes control), vital signs (BP readings), etc., could indicate control. But often an investigator letter from PCP or a thorough medical history is used. - **Medications and life expectancy:** If a patient is on lots of medications or treatments for other diseases, that can cause exclusion. For instance, if someone is on high-dose opioids for chronic pain, they might be excluded as that could affect cognition. Or if they are on warfarin and the trial requires lumbar puncture, maybe that's a problem unless they can hold it. But more systematically, they may just exclude patients with a life-threatening illness like advanced cancer or end-stage organ disease. Life expectancy criteria (like "must have >2 years life expectancy") might be implicitly stated or required by some prevention trials. - **Functional Abilities:** Trials often require that the patient can perform basic personal activities of daily living with minimal assistance. If someone is too functionally impaired (bedridden, etc.), they likely have advanced dementia beyond the trial's scope. Some protocols state as inclusion: "able to perform basic self-care (dressing, eating) with help of caregiver" and as exclusion: "requires skilled nursing care or is institutionalized." Many trials exclude patients who live in nursing homes, partly because it's logistically harder and they tend to be later stage. So usually, participants are community-dwelling or in assisted living at most. EMR might note living situation in social history. Or the caregiver will convey how independent the person is. Some scales like ADCS-ADL (Activities of Daily Living scale) might be used to quantify function at baseline and ensure it's within a certain range. - **Procedural Tolerance:** This can be considered a clinical criterion: patients must be able to undergo study procedures (which often include **MRI scans, PET scans, and possibly lumbar punctures** for CSF). Thus, criteria will exclude: - Those with **severe claustrophobia or inability to lie flat** (since MRI/PET

require that). This is explicitly often listed [35] ("Exclusion: inability to tolerate MRI due to claustrophobia or presence of metal implants..."). - Those with **MRI-incompatible implants** (e.g. pacemakers not MRI-safe) – if MRI is required. If someone has a pacemaker, some trials might allow CT instead if it's just for excluding other pathology, but often they'll exclude because MRI is mandated for uniformity. - Those unwilling to have a **lumbar puncture (LP)** if CSF biomarker is required. Some trials require a baseline CSF sample for biomarkers; if a patient refuses or has a contraindication (like on anticoagulation they can't stop), they can't enroll. Fear of LP is indeed a barrier [39]. - Generally, those unable to complete lengthy **cognitive testing** (which can be 2–3 hours of neuropsychological batteries). If someone has severe sensory deficits (very poor vision or hearing that isn't corrected) – they may be excluded because they can't properly take cognitive tests or see stimuli. This is sometimes explicitly in criteria: "adequate vision and hearing, with aids if needed, to participate in testing." If not stated, it's considered by investigators. EMR might note "macular degeneration, vision 20/200" – that could exclude if tests are visual. Or hearing loss could hamper interviews. - Also, if someone cannot speak or write due to a stroke or other condition, they wouldn't be able to do cognitive tests. So those neuro deficits exclude by default (which ties back to excluding other neurological disease).

These criteria ensure the participant can comply with the study regimen. They aren't exactly in the EMR except if e.g. an MRI report shows artifacts from an implant, or a cardiology note says "pacemaker in place." That can be found (pacemaker often as a Problem or a device noted). Claustrophobia might be found in psych notes or just known by asking the patient. At screening, research staff often specifically ask about "Have you had MRI before? Any issues? Any metal implants? Are you comfortable with needles for LP?" etc. - **Representation in EMR:** Many of these clinical criteria (diagnosis of AD, absence of certain diseases) are reflected in structured diagnoses (Condition list entries) and unstructured notes (neurology consults summarizing the workup). Key cognitive measures (MMSE etc.) are often recorded at least in narrative form. AD patients often have neuropsych evals: those are typically PDF reports or long narrative, but sometimes summary scores might be entered as flowsheet data. The presence of other neurologic conditions like stroke can be found via problem list or past medical history (ICD codes for stroke, or MRI findings in radiology text). Similarly, psych diagnoses might have ICD codes (e.g. Depression F32.9). But sometimes a patient might have significant depression or anxiety not coded, only mentioned. The research team usually does their own screening for depression severity using scales like Geriatric Depression Scale (GDS) or PHQ-9. If the score is above a threshold (like GDS > 6), they might consider that exclusionary if protocol says no significant depression. That GDS would be an Observation not in the routine EMR unless specifically done.

In summary, to find an eligible AD trial participant, one would query: - Age between X and Y (structured), - Diagnosed with Alzheimer's or MCI (structured or via text search for "Alzheimer"), - MMSE in a certain range (if stored, or else a note text search or have to test directly), - Has a listed caregiver (not in EMR, have to ask), - No excluded diagnoses like stroke/Parkinson (structured conditions), - Doesn't have major depression (maybe check problem list, meds like high-dose SSRIs, etc., but likely interview needed), - Doesn't have MRI contraindications (check problem list for pacemaker, maybe see if any MRI had to be aborted in radiology notes, etc. Again often just ask), - Basic labs normal (coming up next).

- **Note on Frequency of criteria:** The analysis by Mitchell et al. [31] highlights age, neurologic, psychiatric criteria as top exclusions. That indicates nearly all trials enforce those in some way. So terms like "age $\geq$ 50," "probable Alzheimer's disease," "MMSE 20-26," "no history of stroke," "no other neurological disease," "no major depression" appear extremely frequently in AD trial protocols.

## Diagnosis Criteria (Alzheimer's)

- **Standard Diagnostic Criteria:** Inclusion typically states that the participant must meet criteria for **Mild Cognitive Impairment (MCI) due to AD or for Dementia due to AD**. For instance, "must meet NIA-AA criteria for probable Alzheimer's disease dementia" or "meet Petersen criteria for amnestic MCI." These criteria entail:
- Gradual onset and progression of memory impairment,
- Objective cognitive deficits on testing,
- Preservation of general function in MCI (for dementia, interference in daily life),
- No other explanation (like normal imaging aside from atrophy, normal labs for other causes, etc.),
- Possibly evidence of AD pathology (in newer criteria, biomarkers).

The protocol often requires that a qualified clinician has made the diagnosis. Many trials have a screening visit where the study doctor confirms the diagnosis by reviewing history, doing a cognitive exam, and confirming the patient fits the diagnosis. They might require a brain MRI to support it (discussed next).

In the EMR, a patient being referred likely already carries a diagnosis of Alzheimer's or MCI from their neurologist or geriatrician. That is a Condition code (like ICD-10 G30.1 Alzheimer's with late onset, or G31.84 Mild cognitive impairment). It might also be in letters and assessments. So structured data: problem list with "Alzheimer's disease" confirms they have the diagnosis. But confirming they meet full criteria often involves seeing the neuropsych test results and excluding other things (which is why they exclude other neuro conditions and require imaging/labs).

- **Exclusion of Other Dementias:** AD trials meticulously **exclude other forms of dementia**:
- **Vascular dementia**: if a patient's cognitive impairment is primarily due to strokes (multi-infarct), they don't fit AD trials. So if MRI shows extensive infarcts and the pattern is more vascular, they'll be excluded.
- **Lewy body dementia**: characterized by prominent visual hallucinations, Parkinsonism early, etc. If a patient likely has Lewy body disease, they would be excluded (some trials specifically exclude those with "core Lewy body features").
- **Frontotemporal dementia (FTD)**: typically presents with behavioral changes or language issues. Trials will exclude patients whose presentation is more consistent with FTD (e.g., early personality change, MRI with frontal lobe atrophy, etc.).
- **Other neurodegenerative diseases**: e.g., Progressive Supranuclear Palsy, Multiple System Atrophy, etc. Not common but they want to exclude any "non-AD neurodegenerative disease."
- **Normal Pressure Hydrocephalus (NPH)**: has cognitive issues with gait disturbance and enlarged ventricles treatable by shunt, so if MRI suggests NPH, those patients aren't typical AD and might be excluded or at least require treating NPH first.
- **Significant cerebrovascular disease**: not just multiple infarcts, but even if MRI shows a lot of white matter lesions (leukoaraiosis), some trials consider that exclusion if severe because it indicates vascular cognitive impairment component.

Many protocols will have an exclusion list: "Any neurological condition that could account for dementia: e.g., Parkinson's disease, stroke, brain tumor, normal pressure hydrocephalus, Huntington's disease, etc." In practice, how they enforce it: through **brain MRI results** (if multiple strokes or tumor, they exclude) and clinical history (if Parkinson's features are present, they might exclude).

EMR contains clues: diagnoses of those conditions in the chart. If someone has an active problem "Parkinson's disease," that's straightforward disqualification. If they have "history of stroke," you'd look

at MRI to see how bad. Many times, potential participants come from memory clinics where they've already weeded out obvious non-AD cases.

- **Biomarker-based Diagnosis (Amyloid PET/CSF):** A big shift in recent AD trials (especially disease-modifying drug trials like anti-amyloid antibodies) is requiring evidence of **amyloid pathology** to ensure the diagnosis of AD is biologically confirmed. That basically means an **amyloid PET scan positive** or **CSF analysis showing low amyloid-beta and high tau**. It was noted that "modern AD trials frequently incorporate biomarker eligibility criteria to ensure AD as the cause" [40]. For example, all Phase 3 trials of amyloid-clearing antibodies (e.g., aducanumab, lecanemab, donanemab) required a positive amyloid PET at screening. If a patient doesn't show amyloid in their brain, they likely don't have AD and are excluded. This criterion wasn't universal historically (older trials of symptomatic drugs often didn't require biomarkers), but now it's common in trials for disease modification. That means the trial has a screening amyloid PET or uses a recent PET if available. The prevalence: among recent trials, a lot in Phase II/III use this; the diversity analysis mentioned that these biomarker requirements can reduce eligibility especially in minority groups [36]. But hard numbers might be like: in one study, among patients screened, up to ~30% turned out amyloid-negative and were screen failures. The trend is requiring it.

Representing amyloid PET: It's an imaging study with a result (positive or negative for amyloid binding). In EMR, an amyloid PET report often gives a qualitative interpretation ("PET scan shows diffuse cortical amyloid deposition consistent with Alzheimer's pathology" or "no significant amyloid uptake – negative scan"). FHIR would have that as a **DiagnosticReport** (category: imaging, modality PET) possibly with coded conclusion SNOMED code ("Amyloid imaging positive"). More likely, just narrative. Some EHRs may bring it in as a PDF from the imaging center. For screening, often the trial sponsor pays for the scan and they get the report. The site uses that to decide eligibility. If using CSF, then a lumbar puncture is done and CSF Abeta42, Tau, phospho-Tau are measured. Those come back as lab values. Many local labs don't do it clinically, so it might be done via research labs. If done clinically, you get numeric values; often they say "CSF Abeta42 < X pg/mL = abnormal, Tau elevated." For eligibility they might define a threshold. If integrated, that'd be Observations (with LOINC codes for CSF beta-amyloid [ for instance, LOINC 97664-3 "Amyloid beta 42/40 ratio in CSF" or similar] and values). If not integrated, the coordinator just sees the lab sheet from the central lab.

So in summary, an increasingly frequent criterion is **"amyloid-positive status confirmed by PET or CSF"**. Without it, the patient is a screen fail in those trials. Some earlier-phase trials and symptomatic drug trials still might not require it, but it's trending toward being standard in disease-modifier trials. We included it under pathology/biomarker domain in Table 3. - **Genetic Criteria:** A few specific contexts: - **APOE ε4 genotype**: Apolipoprotein E (APOE) ε4 allele is a risk factor for AD. Trials usually do not require it for inclusion (since ~40-65% of AD patients have at least one ε4, but not all). However, some trials stratify dosing by APOE4 status because ε4 carriers have higher risk for amyloid-related imaging abnormalities (ARIA) with amyloid antibodies. For example, in aducanumab trials, APOE4 carriers initially had a lower max dose allowed. In some prevention trials (like API trial in homozygous APOE4 individuals), they specifically targeted APOE4 carriers. If so, genotype becomes an inclusion criterion (e.g. "must have 2 copies of APOE4" for that particular trial of young at-risk individuals). That's uncommon except in research specifically on risk or small subgroups. But it's worth noting that many AD trials do perform APOE genotyping at screening (for stratification or safety) even if they don't use it to include/exclude. If a trial had an APOE criterion, the EMR might not have that info unless patient had genetic testing. Usually APOE genotyping is done as part of trial screening if needed (not many people know their APOE from clinical care, since it's not routinely tested because it doesn't change diagnosis or treatment in practice). - **Dominantly Inherited AD mutations:** Some prevention trials (like the DIAN-TU trials) specifically require presence of an autosomal dominant AD gene mutation (APP, PSEN1, PSEN2)

and are only open to those carriers (often identified through the DIAN registry). Those are highly specialized and rare. The EMR would have genetic test results if they underwent testing. But practically, participants for those come pre-selected via research networks. - **Down syndrome:** Some AD trials exclude people with Down syndrome (even though they have high AD risk) because the disease progression and comorbidities differ. Conversely, some specialized trials specifically target AD in Down syndrome. But for general AD trials, having Down syndrome might be an exclusion as an "other neurologic condition" or simply because the trial wasn't designed for that population. EMR would obviously list Down syndrome as a condition if present.

- **Staging and Severity Diagnoses:** They often want a specific range, which we covered in cognitive criteria. Also, they might require that if the patient is in the dementia stage, they have a caregiver rating of how severe – like a specific range on ADCS-ADL or FAQ (Functional Activities Questionnaire). For instance, a mild AD trial might require ADCS-ADL score > 40 (to ensure not too impaired). These are often secondary inclusion checks.

- **Representation and Frequency:** All AD trials revolve around the correct diagnosis and stage, so these terms (MCI, mild AD, probable AD, etc.) are extremely common. The structured representation is partly in diagnosis coding and partly in test results. For example, one could search an EMR for patients with G30.0 (Alzheimer's disease ICD10) and an MMSE score of 24–30 to find MCI/mild AD. That would get candidates. But then exclusion of others is needed.

The analysis by Mitchell et al. found most frequent criteria were age, specified neurologic and psych disorders [31], implying that having the AD diagnosis itself was given (since they specifically looked at criteria that exclude, presumably).

- **Conclusion:** AD trial diagnosis criteria ensure the person truly has AD-type cognitive impairment and not something else. They combine clinical judgment with biomarkers now. In EMR terms, confirming AD often requires looking at multiple data points: problem list (diagnosis), cognitive testing results, MRI findings, specialist notes. That's why many sites have a specific "screening visit" where all this is assembled and reviewed, rather than relying on one click from EMR.

## Prior Treatment Criteria (Alzheimer's)

- **Standard AD Medications:** A key consideration in AD trials is whether participants can be on current standard treatments, namely **acetylcholinesterase inhibitors (AChEIs)** like donepezil (Aricept), rivastigmine, galantamine, and the NMDA antagonist **memantine**. Different trials handle this differently:
- Many trials allow continued use of stable doses of AChEIs and memantine, because it's considered unethical to stop effective therapy (though modest effect) for a long trial. They often *require that doses are stable for a certain period* before entering the trial (e.g. "must have been on a stable dose of donepezil for ≥3 months prior to screening"). This ensures any cognitive changes during trial aren't due to recent med changes.
- Some trials (especially early-phase or certain mechanism studies) might require patients not be on any cognitive enhancers, but that's less common now. If they do, they may have a washout period (like stop donepezil 8 weeks before baseline). But dropping standard of care meds can cause clinical decline and increase variability, so many try to avoid that.
- For example, an inclusion may say "Patients may continue on stable doses of cholinesterase inhibitors and/or memantine if they have been on them ≥12 weeks and will remain at the same dose during the trial." Or conversely, if not on them, some trials allow that too (they won't force start them).

So in summary, a typical criterion: **"If on AD medications (donepezil, etc.), must be on a stable dose for X time. If not on them, patient should not start them during trial."** Or they might encourage being on them but not mandate.

EMR data: medication lists show if they are on donepezil, etc. The research team will verify adherence and stability (often by patient report and med history). If a patient just started or changed dose last month, they might have to wait until stable. If they aren't on any, that's fine unless the trial oddly requires it (rare).

- **Other Medications:** Trials have criteria around **psychoactive medications**:
- They usually allow SSRIs or antidepressants if dose is stable and it's not too sedating. But they might exclude patients on **anticholinergic meds** (like certain bladder meds or some antihistamines) because those can worsen cognition.
- Often they list exclusions like "no use of systemic corticosteroids exceeding X dose" because steroids can affect cognition or underlying conditions.
- **Antipsychotics or benzodiazepines**: Many AD patients might be on a small dose for anxiety or sleep. Trials typically allow low, stable doses at bedtime, but exclude high doses or multiple such meds as that confounds cognitive testing. For instance: "no neuroleptic medication except low-dose quetiapine up to 50mg at bedtime if needed" or "no chronic use of benzodiazepines, except an occasional dose of lorazepam (≤1mg/day) if absolutely necessary." If a patient needs regular xanax or antipsychotic for behavior, they might be excluded as that indicates more severe or unstable condition.
- **Anti-seizure meds**: If taken for seizures, might exclude if seizures uncontrolled. If taken for other reasons (e.g. mood stabilizer), they examine if it has CNS effects.

So medication exclusions revolve around eliminating drugs that could cloud cognitive assessment or indicate another issue. In EMR, active medications can be listed, so the team reviews all meds. If someone is on diphenhydramine nightly for sleep, they might ask them to discontinue and switch to something else because it's anticholinergic. Some trials enforce a washout for certain supplements (like no high-dose vitamin E or omega-3 if they worry it interferes, but that's generally not strict, except certain trials might specify no use of supplements intended to treat memory within X time).

- **Prior Experimental Treatments:** Usually, **no participation in other investigational drug trials recently** (often 30 or 60 days prior). This is a standard criterion across all trials to avoid carry-over effects. If someone was in a different AD trial and took a study drug, they have to wait. EMR wouldn't necessarily know this unless the patient or research records note it. It's typically determined by directly asking the patient and checking any known history.

- **Prior AD Treatments:** There haven't been many approved AD treatments beyond AChEIs and memantine. But if the trial is for an antibody, they might exclude someone who previously received another investigational antibody (to avoid antibody buildup or immune responses). For example, if a patient already got aducanumab on an earlier trial, a new anti-amyloid trial might exclude them. Or require certain time since that exposure. So criteria can say "no prior treatment with any anti-amyloid immunotherapy" or gene therapy, etc. This will become more relevant as new drugs get approved or tried.

- **Medical Procedures Prior:** If a trial requires an invasive procedure like a lumbar puncture, they might exclude those who had complications from a prior LP or major back surgery that precludes LP. But that's more of a technical exclusion.

- **Rehabilitation or Cognitive Therapy:** Some trials might require that patients not be in conflicting cognitive stimulation programs or to keep them constant. Usually not an eligibility criterion but they might ask to keep daily routines consistent.

- **Representation:** The EMR medication list (FHIR **MedicationStatement/MedicationRequest**) provides structured data on what the patient is taking. It's straightforward to see donepezil, memantine, sertraline, etc. Some older patients might have long med lists; research staff go through them to spot any that are disallowed (like say oxybutynin for bladder, which is anticholinergic, they might ask to switch to mirabegron if possible). If a disallowed med cannot be stopped, the patient is out. For structured approach, one could imagine a knowledge base of disallowed drug classes and checking med list against it.

- **Frequency of these criteria:** Very common. The presence/absence of cholinesterase inhibitors, and stable doses, is nearly always addressed in AD trial protocols (noted in inclusion/exclusion section). The same for psychiatric medication limitations. The Mitchell et al. analysis didn't explicitly mention these in abstract, but they would fall under "procedural" or "medical" categories.

- **Lifestyle:** Also, AD trials might exclude those with **substance abuse**. They often require no alcohol abuse (like "no more than 2 drinks/day" or if someone meets criteria for alcohol use disorder, exclude). Tobacco use isn't always exclusion unless heavy and causing health issues, but sometimes if a PET ligand is used, heavy smoking could affect it (rare scenario). So usually not a strict criteria except "no drug abuse."

- **Caregiver Training**: If a trial involves the caregiver in delivering an intervention, they might need certain caregiver criteria (like "caregiver is able to speak English and has at least 8th grade education" if they must do tasks). This is trial-specific, not general.

## Imaging Criteria (Alzheimer's)

- **Baseline MRI to Rule Out Other Pathology:** Most AD trials require a **brain MRI at screening** for diagnostic confirmation. The MRI is checked for:
- **No large strokes or hemorrhages:** Some small vessel disease is expected in older folks, but they often set a threshold. E.g. exclude if there is any cortical infarct >1cm or more than 2 lacunar infarcts, etc. Or if there's any sign of prior intracerebral hemorrhage bigger than microbleeds. This is to avoid vascular dementia or risk of hemorrhage if giving an amyloid drug (which can cause ARIA-H, microhemorrhages).
- **No structural abnormalities** like a brain tumor, significant normal pressure hydrocephalus (marked ventriculomegaly with gait disturbance), or subdural hematoma, etc. These things, if present, suggest a non-AD cause for dementia or a treatable condition.
- **Appropriate level of atrophy and white matter changes for age:** They won't exclude for just having atrophy, since AD has hippocampal/ cortical atrophy. But extreme atrophy might indicate advanced stage beyond the trial's target (though usually stage is determined clinically, not by volume).
- **No severe microvascular ischemic changes**: Many MRI reports give a Fazekas score for white matter lesions. Trials sometimes exclude patients with "severe" white matter disease (since that correlates with vascular cognitive impairment). They might allow mild to moderate changes. This is often a judgment call by the investigator or radiologist. Some protocols define it, e.g., exclude if >1.5 on a 0-3 scale of leukoaraiosis or if confluent white matter disease.
- **No more than X microhemorrhages**: In trials of amyloid-lowering antibodies, this became crucial. Amyloid-related imaging abnormalities with hemorrhage (ARIA-H) risk is higher if a

patient already has microbleeds. E.g., some protocols originally excluded if >2 microhemorrhages on gradient-echo MRI [41] [42], but then an expert group said up to 4 is okay [43]. So now often "exclude if more than 4 microhemorrhages or any superficial siderosis on MRI" because those indicate risk for ARIA-E/H events. So reading the GRE or SWI MRI sequences is part of screening for amyloid trials.

- **If contrast MRI is used**: ensure no unexpected enhancement (like inflammation or tumor).

So effectively, **MRI exclusion criteria** are a big component of screening. They ensure pure AD pathology and safety for anti-amyloid treatments.

EMR wise: the MRI is done usually as part of trial screening (could be billed to sponsor, but often goes in EMR as a research MRI). The radiologist will produce a report. The site PI also reviews images for trial-specific aspects (like counting microbleeds). The formal radiology **DiagnosticReport** (in FHIR or PACS) contains descriptions: e.g. "Moderate white matter T2 hyperintensities, a 5mm chronic lacunar infarct in left thalamus, and 3 microhemorrhages in frontal lobes. No mass or hydrocephalus." Based on that, the site determines eligibility. The structured representation is limited (a radiology report could be partially structured, but usually free text). The presence of specific findings might be encoded in some template (some MRI centers use a standardized dementia MRI report including an "ARWMC score" for white matter changes, etc.). But typically, the text suffices. - **PET Imaging for Pathology or Metabolism:** We already discussed **amyloid PET** as a biomarker for inclusion. That is imaging but could be considered pathology evidence. Similarly, some trials might use **FDG-PET** to ensure an AD-like metabolic pattern (rare nowadays, used more in differential diagnosis research). Or **tau PET** (newer trials might even require tau PET to confirm enough pathology, though so far tau PET is more often an outcome measure than eligibility, except perhaps requiring "no off-target binding" stuff, which is not likely an exclusion except in research contexts). If any PET is required just for screening (like amyloid PET), then the criterion is basically "amyloid PET must be positive." If negative, exclude. Already covered that. Tau PET might become a criterion in future (like require tau pathology in a certain range if testing an anti-tau drug).

- **No MRI Contraindications:** As mentioned under clinical, if a patient has a pacemaker or metal implants that prevent MRI, they're typically excluded unless the protocol has an alternative (some allow CT if absolutely can't do MRI, but that is rare, CT is not as sensitive for microbleeds or small strokes). So the presence of any MR contraindication (like ferromagnetic aneurysm clip, severe claustrophobia not alleviated by sedation) is an exclusion. EMR often lists in the allergy field or problem list something like "MRI Safety: Pacemaker" or radiology records might note "MRI not done due to pacemaker." That is considered.

- **Follow-up MRIs:** Many AD trials do periodic MRIs to monitor ARIA (for amyloid drugs). If a patient shows ARIA-E (edema) or hemorrhages, they might be temporarily taken off drug. But that's not initial eligibility, it's on-study monitoring. However, initial presence of microbleeds influences risk. So initial imaging criteria heavily influences who gets in to mitigate that.

- **Volumetric MRI or Atrophy Requirements:** Usually they don't require a certain atrophy level as a criterion (that might bias sample), but sometimes in early trials they wanted to ensure patients had some hippocampal atrophy to enrich for AD. In modern practice, biomarker (amyloid PET) is used for enrichment rather than relying on atrophy. So we can skip that.

- **Imaging for Excluding Other Dementias:** We basically covered: if MRI shows things like cortical Lewy bodies can't be directly seen, but if MRI is normal and cognitive profile suggests FTD, they might find frontal predominant atrophy and exclude if not typical of AD. But mostly, MRI's role is to exclude vascular lesions and other pathologies.

- **Representation Summary:** Imaging eligibility in AD is largely a manual but systematic review of MRI findings. Radiology **DiagnosticReport** text holds the information, often with sections like "Findings: moderate small vessel ischemic changes, no acute infarct, etc." and "Impression: Findings consistent with age, no evidence of alternate pathology to explain dementia." Some trials require sending the MRI to a central reader to confirm eligibility (especially if measuring microbleeds, they might have central radiologists count them reliably). But initial local check is done too. There is no quick structured data for "number of microhemorrhages" except if radiologist explicitly states it. Maybe counted in text. A future FHIR extension could list findings, but not in standard use yet.

- **Procedural imaging tolerance:** We touched on MRI tolerance (no claustrophobia). They may allow a trial if patient can get an MRI under sedation; some protocols allow a mild sedative before MRI if anxiety is an issue, as long as it doesn't significantly alter cognitive state at baseline (they usually do MRI on a separate day from cognitive testing anyway). But if someone outright refuses MRI or couldn't complete one, they cannot be enrolled. That's a screen fail reason often recorded ("screen failure due to inability to tolerate MRI").

- **Screening vs Ongoing Imaging:** AD trials typically have at least 1 MRI at screening and maybe at end (to check ARIA or measure brain volume changes). Some have periodic if needed for safety (like amyloid drug trials often do MRI at 3, 6, 9 months to monitor ARIA-E). Those are not criteria but safety monitoring.

## Pathology/Biomarker Criteria (Alzheimer's)

- **Amyloid Biomarker Positivity:** This is worth re-emphasizing as a relatively new *inclusion* criterion in many trials. As noted, the **amyloid PET** or **CSF amyloid** result must indicate AD pathology. Trials like aducanumab's involved screening out ~30% of clinically diagnosed AD patients who turned out amyloid-negative. That shows why it's critical for trial success (to avoid including non-AD patients who wouldn't respond and would dilute results). So nowadays:
- **Anti-amyloid drug trials** – almost all require amyloid positivity (e.g., "evidence of brain amyloid pathology on PET or CSF").
- **Anti-tau drug trials** – some of these might also require amyloid positivity because they want to ensure it's AD, or some might even require tau PET (though tau PET is not as widely available; some trials might just ensure amyloid positive, reasoning if they have amyloid, they likely have tau).
- **Symptomatic trials (like cognitive enhancers)** – historically did not require biomarkers, but new ones often still prefer it, to make sure they have AD and not another dementia type.
- For **MCI trials**, since MCI is heterogeneous, confirming amyloid is especially important (to distinguish prodromal AD from other mild impairments). So most MCI trials for disease-modifying agents require an amyloid positive scan or CSF.

If a patient's EMR already has an amyloid PET (like if done clinically or in a prior study), that could be used if recent. If not, the trial does one. So from an EMR perspective, it may not be present beforehand unless they were in a diagnostic study (some memory clinics order amyloid PET if available and insurance covers, but that's still not routine for every patient).

- **CSF Biomarkers:** If PET isn't available or as alternative, **CSF analysis of Aβ42, total tau, phospho-tau** can be used. Criteria might set a threshold like "CSF Aβ42 < 500 pg/mL or an Aβ42:tau ratio below X, consistent with AD." The specifics vary by assay. It's either/or with PET typically. Some smaller academic trials might only use CSF if they don't have PET access. The presence of this criterion means the site must do a lumbar puncture or have results from one.

Clinically, sometimes lumbar punctures are done if diagnosis is uncertain. FHIR Observations would capture these lab results with codes for those analytes. E.g. LOINC 2965-2 for CSF protein tau, etc., with a value. Trials might define positivity by some lab's reference: e.g. "CSF Aβ42 < 400 pg/mL is positive." These numeric criteria are evaluated by the lab that runs it (like the local or central lab returns a result flagged abnormal).

- **Tau PET:** Not standard as an entry criterion as of now, but some trials might require a *minimum level of tau pathology*. Donanemab's trial, interestingly, **excluded** patients with very high tau PET uptake (since they hypothesized those might be too advanced to benefit). They used tau PET to select an intermediate tau burden group. That's unusual and specific: they wanted to exclude either very low tau (which would be earlier than AD, maybe just amyloid positive cognitively normal?) and very high tau (very advanced). So in that case, tau PET was both done and used for excluding either extreme. That might become more common – tailoring who is likely to respond by pathology stage. So one could see criteria like "Tau PET shows Braak stage I–III equivalent binding (moderate), excluding those with high global tau burden." Representing that in EMR is tricky – tau PET results are usually given as SUVR values or regional scores, and currently that's research. But if such criterion exists, they usually do it centrally or at least with a specific analysis pipeline. In the future, if tau PET is widely used, it could appear in diagnostic records (like "tau PET positive in temporal lobe" etc.). But for now, think of it as a specialized criterion in few trials (donanemab's being a prominent example).

- **Genetic Biomarkers:** We discussed **APOE genotype** – rarely a requirement except specialized trials (like some prevention trial targeted at APOE4 homozygotes had that as inclusion). Usually, genotype is not used to select in late-onset AD trials because it's not necessary (and excluding APOE4 negatives would reduce sample and maybe skew results).

- **Neurodegeneration Markers:** Some trials might require evidence of neurodegeneration aside from amyloid – like MRI atrophy or CSF tau. In NIA-AA research framework, "ATN" (amyloid, tau, neurodegeneration) markers are considered. For example, a trial might want to ensure patient has both amyloid and elevated tau (which often correlates with disease stage). But practically, requiring amyloid positive is already narrowing to AD.

- **Blood Biomarkers:** Very recent development: blood tests for AD (plasma Aβ, phospho-tau) are emerging. Possibly in the near future, trials may use a blood test to pre-screen or even as part of eligibility. E.g., a high plasma p-tau could enrich for AD. As of 2025, I'm not aware of trials that use blood biomarkers as formal inclusion criteria (most still rely on PET or CSF which are more proven). But sites sometimes pre-screen with blood tests to decide who to send for PET (cheaper to do blood first). That might not be in protocol, just site practice. If it were in criteria, one could see "plasma p-tau > threshold" needed, but I think not yet standard. If it comes, that would be just another lab Observation to check.

- **"Biomarker-negative" exclusion:** Another way to phrase: effectively, "Exclude patients who lack evidence of AD pathology." Many symptomatic trials in the past didn't check, and likely included some non-AD – which might partly explain trial failures. So the field moved to requiring pathology evidence.

- **Exclusion of other pathology via biomarkers:** They might exclude if CSF shows infection or something (but that's rare). They often test syphilis serology or HIV to exclude those etiologies of dementia (which is more lab domain, but also part of diagnostic work-up).

- **Representation in EMR:** Many of these pathology biomarkers (amyloid PET, CSF results, genetic tests) are not standard clinical tests historically, so they might not appear in an EMR unless done as part of clinical eval or previous research. If done clinically: e.g., a patient had a clinical amyloid PET (which Medicare covers in limited cases as of 2025 via certain programs), the result might be in EMR. Genetic testing for AD genes or APOE if done clinically (rarely done outside research or direct-to-consumer) could be in a genetics clinic note or in an uploaded result.

But increasingly, as disease-modifying therapies like lecanemab (approved in 2023) come into practice, confirming amyloid positivity is needed even clinically to qualify for those treatments. So hospitals are starting to get more amyloid PETs in EMR for treatment decisions. That might ironically make it easier to find candidates for trials (e.g., "this patient is amyloid PET positive but can't get drug due to something, maybe they can join trial X").

- **Prevalence in Protocols:** Historically not all trials did this, but now the major trials do. For instance, NEJM Evidence paper [36] suggests modern trials often incorporate such criteria. Possibly by 2024, a majority of Phase 3 AD trials had amyloid as required. The earlier-cited stat of 65% neurologic and 61% psych is about excluding other conditions, which indirectly implies doing biomarkers to confirm AD (not directly given as number, but one can infer that at least large chunk do).

- **Conclusion:** The pathology/biomarker domain in AD trials is about verifying that the target pathology (amyloid, etc.) is present and others are absent. It's becoming a cornerstone of trial enrollment to reduce misdiagnosis.

## Laboratory Criteria (Alzheimer's)

- **Laboratory Tests to Exclude Other Causes of Dementia:** In standard AD diagnostic practice, one does lab tests to rule out **reversible causes of cognitive impairment** such as:
- **Vitamin B12 deficiency:** Low B12 can cause cognitive issues; if found, patient should be treated and then if cognition improves it wasn't AD, if not and other evidence suggests AD, it's coexisting.
- **Thyroid dysfunction:** Hypothyroidism especially can cause memory problems; hyperthyroid can cause confusion too. So TSH (thyroid-stimulating hormone) is usually tested. If abnormal, the patient is treated and not included until stable. Most trials say exclude if thyroid disease is uncontrolled. If hypothyroid but on medication with normal TSH, that's fine.
- **Folate deficiency:** sometimes tested with B12.
- **RPR (syphilis serology):** Some guidelines recommend it to rule out neurosyphilis (rare but treatable cause). Many memory clinics do it once. Trials often require a negative syphilis test in screening or history.
- **HIV test:** HIV-associated neurocognitive disorder could mimic dementia, and active HIV can predispose to cognitive issues. Many trials exclude HIV-positive individuals (especially if untreated) or require stable treated status. Historically, like in cancer, HIV patients were often excluded. Now if someone is well-controlled and it's believed they have AD, some trials might allow it if on stable ART and no AIDS dementia, but often they just exclude to be safe.

So an inclusion might indirectly be: "patient's workup for dementia is negative for alternative causes (normal B12, normal TSH, etc.)." Or listed as exclusion: "abnormal B12 or TSH that could explain cognitive impairment" (unless corrected).

In EMR, these lab results are normally present in diagnostic workup. If not, the trial will do them at screening. They are Observations: e.g. B12 level (LOINC 2132-9), TSH (LOINC 3016-3). The site ensures

those are within normal (or if not, that it's been addressed; e.g. if TSH was high but now on synthroid with normal TSH, okay).

- **General Health Labs:** AD patients being older often have comorbid conditions, so trials do basic labs to ensure no major organ failure:
- **CBC:** to make sure no severe anemia (which can cause confusion) or infection.
- **Liver function tests:** ensure no hepatic encephalopathy potential or active hepatitis. If someone has AST/ALT 3-4x ULN from, say, NASH, they might still allow if stable, but some protocols might exclude if, for example, ALT > 3x ULN just for safety with experimental drug metabolism. The threshold may not be as strictly enforced as in oncology, but it's looked at.
- **Renal function:** ensure not on end-stage (someone on dialysis might be excluded, partly because they have limited lifespan and difficulty coming to visits). But moderate CKD is common in elderly; usually allowed if stable and not affecting cognition (though severe electrolyte issues could cause confusion).
- **Glucose/HbA1c:** if uncontrolled diabetes (very high sugar) can cause cognitive issues. They may require that it's under control (HbA1c below certain threshold, e.g. <10%). Many older AD patients have type 2 diabetes; they aren't excluded if controlled, but if someone's A1c is 13% (poorly controlled), that's an exclusion until improved, because hyperglycemia can cause fatigue, cognitive clouding, and risk of acute issues.
- **Electrolytes:** no major imbalances. E.g., chronic hyponatremia (low sodium) can cause cognitive problems. So if sodium is, say, 128 mmol/L (mild hyponatremia), they might hold off or treat that. Trials often check it and if severe, exclude until fixed.

But these labs are more about safety and excluding someone whose mental status might be affected by another condition. The criteria might not list them explicitly except in broad terms ("no clinically significant laboratory abnormality"). If the investigational drug has particular organ concerns (like if metabolized by liver, they might have a specific threshold for LFTs like 2x ULN, but many AD drugs are antibodies or CNS-acting small molecules not particularly liver-toxic, though they still watch LFTs).

For example, a protocol might have an exclusion: "AST or ALT > 3 × ULN; total bilirubin > 1.5 × ULN; serum creatinine > 2.0 mg/dL or clinically significant renal impairment," etc. Many do include such standard cut-offs, similar to oncology but maybe slightly more lenient. Because AD patients often are physically healthier (somewhat) than cancer patients in trials, these lab exclusions typically don't filter out a huge fraction, but they do catch a few with e.g. unknown hepatitis or very abnormal labs.

- **Preventive or Early-stage Trials Labs:** If a trial is in very early (like cognitively normal but with amyloid or at-risk carriers), they might have stricter lab criteria to ensure those individuals are otherwise healthy. But those populations tend to be younger and healthier anyway.

- **Pregnancy:** It's unlikely for typical late-onset AD patients to be pregnant (since most women are postmenopausal by trial inclusion age, e.g. 50+). But if a trial includes younger presymptomatic individuals (like in a dominantly inherited AD gene carrier trial, participants could be 30s or 40s, some female could be of childbearing potential), then pregnancy testing and contraception would be required. For example, DIAN trials or prevention trials could have that. So just as with others: if applicable, a negative pregnancy test at baseline and agreement to use contraception if fertile. This is rare in AD context but not impossible. The standard approach stands: do HCG test (Observation), ensure negative. It's not needed for most typical AD dementia trials with older participants.

- **Drug screening:** Some trials might do a urine drug screen for illicit drugs, to exclude those regularly using e.g. cannabis or stimulants that might affect cognition or indicate unstable

lifestyle. If they do, that's a lab test as well (qualitative). Particularly if they suspect a patient might be abusing substances, they'd test at screening. Many protocols include "no drug abuse" which might implicitly allow testing for it. It's more common in younger populations (like MCI in 50s).

- **Lab Data in EMR:** All these lab results come from standard lab panels:

- B12, TSH (and possibly free T4) – to rule out deficiency/hypothyroid.
- CMP (comprehensive metabolic panel) – gives electrolytes, glucose, liver enzymes, creatinine.
- CBC – for anemia or infection signs.
- Possibly RPR (rapid plasma reagin for syphilis) – usually done once clinically; if not, they'd do it.
- HIV test – ethically, patient should be counseled. If positive, they might exclude or only include if well-controlled and not cognitive from HIV. Many would exclude because they might not want to deal with confound and ARV interactions.
- Urinalysis – maybe check no UTI causing confusion at baseline (if a UTI, treat then maybe rescreen).

These are all Observations in the lab category, which can be retrieved from EMR easily. Usually they require these to be done within a certain window (like within 30 days of baseline).

- **Prevalence in Trials:** The Mitchell et al. paper didn't highlight labs specifically aside from mentioning disparities e.g. some criteria that can exclude minorities more (like requiring normal lab values might exclude some with comorbidities that are more prevalent in some populations). But overall, I would say nearly all trials have these lab check criteria. Historically, many patients aren't excluded due to labs except maybe B12 or thyroid, which if low, can be supplemented and then patient reconsidered later.

- **General vs Strict:** AD trials often have general phrasing like "laboratory values must be within normal limits or if outside, not clinically significant as determined by investigator." This gives flexibility – unlike oncology where strict cut-offs are used. For instance, if an AD trial sees ALT is 1.8x ULN in an older patient due to fatty liver, an investigator might judge it's fine (especially if the drug isn't hepatotoxic). If it's 5x ULN, probably exclude until resolved. So criteria might not enumerate each threshold but rely on PI judgment. We put some thresholds in table as typical, but actual protocols might just say "no clinically significant hepatic/renal dysfunction." That is subjective – one PI might allow a bit higher values, another might not, though usually sponsors give some guidance on what "clinically significant" means or in investigator meeting they clarify.

- **Therefore:** In practice, trial site will look at labs – if something glaring (like creatinine clearance 30 ml/min – that's quite low; they might exclude for that because long trial might be tough if kidney failing, also some drugs might accumulate), or hemoglobin 9 – maybe exclude until treated (maybe the patient had iron deficiency treatable).

- **Representing Decisions:** Some these things are recorded as "screen failure because of labs" in trial logs. EMR might not explicitly say "excluded because B12 was low" – instead they treat B12 and possibly enroll later if patient improves. If B12 deficiency was found, ethically the doctor treats it and might hold off trial until re-evaluation to ensure cognitive issues remain (if they fully reverse after B12, maybe it wasn't AD).

In summary, AD trial eligibility has a heavy focus on age, cognitive testing, the diagnosis and exclusion of other causes (which is supported by labs and imaging), and ensuring a caregiver is present. Frequency and emphasis differ from oncology: instead of performance status and organ function being

the main limiting factors, it's cognitive stage and absence of other brain diseases (often confirmed by biomarkers) that are paramount. However, the underlying approach of using both structured and unstructured data to verify criteria is common to both domains, as we've elaborated.

## Summary and Data Representation

Across these three indications – lung cancer, breast cancer, and Alzheimer's disease – we see that eligibility criteria consistently cover multiple domains (demographic factors, clinical performance/ comorbidities, diagnosis verification, prior treatments, imaging findings, pathology/biomarkers, and lab results). These criteria serve to define a trial's target population and ensure safety. Table 1, Table 2, and Table 3 summarized the key criteria and their prevalence in each domain for each disease.

From an informatics perspective, many of these criteria have corresponding data points in hospital EMRs: - **Demographics** like age and sex are structured in patient records (FHIR Patient resource), while factors like caregiver availability or language proficiency may only appear in unstructured notes or intake forms. - **Clinical status** items such as ECOG performance (in oncology) or cognitive test scores like MMSE (in AD) can be stored as structured **Observations** with standardized codes [10] [33], but often they are additionally documented in narrative form. Comorbid diagnoses are listed as **Conditions** (problems) in structured form, and also described in history and consult notes. For example, an exclusion like "no prior stroke" would be checked by absence of a stroke ICD code and by reviewing neurology notes. - **Diagnosis confirmation** is captured by the presence of a coded diagnosis (Condition with a confirmed status for cancer or AD) [44] [17] and by the supporting evidence (pathology reports for cancer, cognitive assessments and criteria fulfillment for AD). FHIR's **Condition** resource holds diagnosis details (like "Invasive ductal carcinoma of breast, confirmed" or "Alzheimer's disease, probable"). The **verificationStatus** in Condition can indicate confirmed diagnosis [17]. Unstructured pathology narratives and neurologist's reports complement this by giving details like subtype (ER-positive, etc.) or AD subtype (with or without vascular component). - **Prior treatments** are logged in EMRs as medication administrations, prescriptions, and procedure records. In FHIR, past chemotherapy or targeted therapy can be represented as **MedicationAdministration/MedicationStatement** resources (with drug codes), and surgeries or radiation as **Procedure** resources. However, often a summary of past treatments is written in notes, which the trial team uses to gauge eligibility (e.g. how many lines of chemo, what drugs). Tools can query medication histories to see if a patient has taken a disallowed drug (like a prior EGFR inhibitor) by scanning for that drug name or code. Yet, human review remains crucial to interpret complex histories. - **Imaging results** are structured in EMRs as well – each imaging test yields a **DiagnosticReport** with findings [45] [18]. The content of those reports, such as "no brain metastases seen" or "white matter changes present," is typically in free text [19]. FHIR's **ImagingStudy** resource lists the images and metadata, and the **DiagnosticReport** for imaging contains impressions and can reference specific key images [29]. For trial criteria, relevant imaging info (like presence/absence of metastases, number of microbleeds, etc.) is extracted from these reports. In some cases, structured indicators exist (for example, certain PACS systems might allow tagging "metastasis = yes" in a template), but generally natural language processing or manual reading is needed. Ensuring imaging criteria are met often requires physician interpretation, which is then documented (e.g., an oncologist writes "MRI brain reviewed – no metastases, meets criteria"). - **Pathology and biomarker data** are partly structured: labs and pathology departments often produce **DiagnosticReport** entries for pathology with coded diagnoses and may include structured fields for receptors or mutations. For instance, the FHIR **DiagnosticReport** for a pathology might have category "pathology" and include observations for ER positivity or gene mutations [18] [19]. Genetic and molecular test results can be represented as Observations (e.g. genetic variant found) or as part of a **Genomics report**. In practice, trial screeners often rely on the textual pathology report plus any attached lab reports for mutations or biomarkers. If an EHR is advanced, these results might be searchable via codes (e.g., find all patients with SNOMED code for EGFR mutation). If not, one might search within documents. For AD, biomarkers

like amyloid PET results are in imaging reports (text indicating positive/negative) and CSF results come as lab reports. Those could be structured (CSF Abeta value as a numeric Observation). The **US Core DiagnosticReport Profile** and lab Observations make it possible to share and query such results [22]. For example, a system could query "Observation?patient=X&code=LOINC:95209-3" (if that were a code for amyloid PET result) to see if amyloid was positive. However, implementation of such coding is still emerging. - **Laboratory values** like blood counts, liver enzymes, creatinine, etc., are among the easiest to match with criteria because they are numeric and standard. In FHIR, each is an **Observation** typically also summarized by a **DiagnosticReport** for a panel [22]. These come with reference ranges and can be programmatically checked: e.g., Observation code 1988-5 (ALT) value 100 U/L vs reference 0–40 means 2.5× ULN, which might violate a criterion. The use of standard units and reference ranges in lab data allows automated comparisons if one sets up those rules [24]. Tools could automatically flag a patient whose labs don't meet criteria (like GFR too low, etc.). In practice, research staff review the latest lab results from the EMR or order a fresh screen lab panel, and then make the judgment. Many trial management systems do incorporate lab integration to mark eligibility status (often green/red highlighting if values out of allowed range).

In summary, **high-quality data from EMRs can greatly facilitate trial eligibility screening**. Structured EMR data aligned with FHIR standards can be queried to filter candidates (for example, finding all patients with a Condition "Alzheimer's disease" and Observation MMSE 20-26, etc.). At the same time, **human interpretation of unstructured notes and reports remains essential** for nuanced criteria (like judging whether comorbidities are "well-controlled" or whether an MRI's subtle findings exclude the patient).

Each of the frequent eligibility criteria we identified (age limit, performance status, specific diagnosis confirmation, prior therapy count, absence of brain metastases, presence of HER2 overexpression, amyloid PET positivity, lab thresholds, etc.) has a footprint in the clinical data: - If data is **structured**, it will be found in resources such as Patient, Condition, Observation, DiagnosticReport, Procedure, MedicationStatement. - If **unstructured**, it will appear in textual reports (radiology, pathology) or clinical narratives (history & physical, progress notes, consults).

The table below gives a few examples of how a specific eligibility criterion is represented in structured vs unstructured form:

| Eligibility Criterion | Structured EMR Data | Unstructured Data |
|---|---|---|
| Age ≥ 18 (Lung/ Breast/AD) | Patient.birthDate (calculate age) – stored in EHR; FHIR Patient resource [1]. | Mentioned in notes ("55-year-old male…"). Rarely an issue after initial automated check. |
| ECOG Performance Status 0–1 (Oncology) | Observation with code for ECOG score (e.g., LOINC 89247-1) and value 0 or 1 [10]. Could be in a flowsheet or mCODE profile. | Clinic note "ECOG PS 1" [9] or "Karnofsky 80%" in oncology assessments. These notes confirm functional status. |
| Histologic confirmation of cancer | Pathology DiagnosticReport with SNOMED diagnosis ("adenocarcinoma of lung") [18]; Condition resource with code for that cancer and verificationStatus=confirmed [17]. | Pathology report text ("invasive ductal carcinoma, ER 8/8, HER2 3+"). Oncologist note: "biopsy-proven NSCLC." |

| Eligibility Criterion | Structured EMR Data | Unstructured Data |
|---|---|---|
| Specific cancer subtype (e.g. HER2+) | Observation for HER2 status (LOINC code for IHC result, value "Positive (3+)") or as a component in pathology report; possibly a flag in Oncologist's database. | Pathology report narrative ("HER2: IHC 3+") or separate lab report [27]. Oncologist might write "HER2-positive metastatic breast ca." |
| Brain metastases exclusion | ImagingStudy/DiagnosticReport for brain MRI; no structured field "metastasis present" unless NLP derived. Possibly a coded problem "Brain metastasis" if known. | Radiologist's impression: "No evidence of intracranial metastasis" or "2 metastases in cerebellum." [12] Oncologist's note might summarize "Brain MRI clear." |
| Measurable disease by RECIST | Not directly structured; target lesion measurements could be Observations if entered (each lesion size). Otherwise requires reading radiology. | Radiology report: lists lesions with size. Or trial screening form listing target lesions. If none >1 cm, not eligible (noted by clinician). |
| No prior chemotherapy (in 1L trial) | MedicationAdministration/Statement – no records of chemo drugs (e.g., no drug code for paclitaxel, etc.). Could search pharmacy history. | Oncology notes saying "treatment-naïve" or listing no prior therapies. If prior adjuvant chemo, might be documented ("had AC-T in 2019"). |
| ≤1 prior metastatic regimen (oncology) | Medication history – count of distinct chemo lines. Not trivial to compute from raw data; needs logic to group by lines of therapy. | Doctor's summary: "She received first-line letrozole, now progressing – one prior line." This text is used to confirm eligibility for second-line trial. |
| Lab: ANC ≥ 1500/μL | Lab Observation (LOINC 26499-4 or a complete blood count panel) with ANC result, e.g. 1.8 x10^9/L. Marked as normal or low by system. Easily queryable [22]. | Lab report printout or screenshot in chart, or note "ANC 1800, adequate." If low, might say "neutropenia present, patient ineligible". |
| Lab: AST ≤ 2.5×ULN, bilirubin ≤1.5×ULN | Lab Observations (LOINC codes for AST, ALT, bilirubin) with values and reference ranges. FHIR Observation.referenceRange holds ULN for context [22] [24]. Calculation (value/ULN) can be automated. | Lab report text: "AST 60 (H) [normal 0-40]" or physician note "LFTs less than 2× ULN." Rarely written explicitly in notes unless assessing eligibility. |
| Lab: creatinine clearance ≥ 60 mL/min | Possibly an Observation for eGFR (if lab auto-calculates). Otherwise compute from serum creatinine (Observation) + demographics. Structured if EHR provides eGFR. | Note might mention "CrCl ~65 mL/min by Cockcroft-Gault." Often done behind scenes by coordinator using lab values. |

| Eligibility Criterion | Structured EMR Data | Unstructured Data |
|---|---|---|
| Negative HIV/ HBV/HCV serology | Lab Observations: e.g. HIV antibody test (result negative), Hepatitis B surface antigen (neg), Hep C PCR (neg). These are discrete tests. | Lab result forms or a note "HIV, hepatitis panel negative." If positive, an ID consult note might exist – that would flag exclusion. |
| Negative pregnancy test | Lab Observation: hCG qualitative or quantitative (e.g. LOINC 803-1 urine pregnancy test). Value = negative [34]. | Lab report or nurse note "urine preg test negative on admission." Usually treated as routine pre-trial checklist rather than described in narrative. |
| MMSE score 20–26 (mild AD) | Cognitive assessment could be an Observation (some EHRs have MMSE as a vital sign). If not, stored in research database. Could use LOINC 35619-0 "MMSE score total". | Neurologist note: "MMSE 23/30" [33]. Or psych report listing scores. Trial screening sheet would record the score as well for eligibility. |
| CDR = 0.5 or 1 (very mild/mild dementia) | Possibly in neuro clinic flowsheet or research form (CDR usually not in standard EHR fields). Could be an Observation if recorded. | Memory clinic note: "CDR = 1 (mild dementia)." If not documented, the study team administers the CDR interview at screening. |
| No major depression (in AD trial) | Depression screening scale Observation (e.g. Geriatric Depression Scale score) if done; active problem list – no major depressive disorder ICD code, or if present, note it's in remission. | Psych notes: "Depression in remission on sertraline." Trial might require a scale score: e.g. "GDS = 3, no significant depression." Otherwise, exclusion: "Patient has active depression, ineligible." |
| No other neurologic disease (AD trial) | Problem list review – no ICD codes for Parkinson's, stroke, etc. Radiology – no significant stroke on MRI (see imaging). | Neurology consult: "No history of stroke or other neuro disorders." MRI report mentions any infarcts (if yes, might exclude depending on size/ number). Documented in screening note if any exclusionary lesions. |
| Amyloid PET positive (AD trial) | PET scan report – might have a structured summary (some reports say "Image Interpretation: Positive for amyloid"). Could be encoded as an Observation (e.g. SNOMED code for abnormal tracer uptake) but typically textual. | PET report text: "Diffuse cortical uptake of Amyvid, consistent with amyloid deposition – Positive scan." If negative, report says no significant uptake. Trial site records result ("amyloid PET positive on 1/5/25" or if negative, they're a screen fail). |

| Eligibility Criterion | Structured EMR Data | Unstructured Data |
|---|---|---|
| CSF AD biomarkers (AD trial) | Lab results from CSF: e.g. Observation for Aβ42 (with low value flagged), tau, p-tau. These come from specialized lab. Structured if integrated (LOINC codes exist). | Lab report from CSF analysis: "Aβ42 200 pg/mL (low), Tau 800 pg/mL (high) – consistent with AD." Usually a PDF or printout in chart. The PI writes in screening note "CSF profile consistent with AD, meets biomarker criterion." |
| APOE genotype (if required) | Genetic test result – could be stored as a Molecular Sequence or Observation (APOE genotype test with result E3/E4, etc.). Possibly found in genetics consult note. | Lab report or note: "APOE genotyping: ε3/ε4." If trial required ε4/ε4, and patient is ε3/ε4, then ineligibility noted. Often this is done by trial's lab rather than pre-existing in EMR. |

This comparison illustrates how the **FHIR standard can represent much of the key data** (diagnoses, observations, lab results, reports) that corresponds to eligibility criteria. In a well-structured system, one could programmatically screen some criteria: e.g., fetch Patient age, query Conditions for exclusions (like metastases or comorbidities), retrieve latest labs, etc., to get a shortlist of potentially eligible patients [22] [19]. However, full determination still involves reading unstructured content, especially for nuanced clinical judgments and imaging interpretation.

Finally, to **tie back to frequency analytics**: The common criteria we identified are not just anecdotal – they have been quantified in studies: - In breast cancer trials, upper age limits were present in 39%, strict lab/comorbidity exclusions in 77%, ECOG PS limits in 69% [1]. Bone marrow, liver, renal function criteria appeared in 84%, 51%, 42% of trials respectively [3]. - In lung cancer trials, performance status and organ labs are nearly universal, with one review noting 90% excluded patients with moderate renal impairment [8] and a majority historically excluded untreated brain metastases (only ~17% allowed them by 2022) [6]. - In Alzheimer's trials, **87% had age cutoffs**, 65% specified neurologic exclusions, 61% psychiatric exclusions [31]. Virtually 100% had cognitive score criteria. Caregiver requirement is effectively 100% in AD dementia trials as well [32]. Modern trials increasingly require amyloid positivity (which was not quantified in that study but is evident in recent phase 3 protocols).

These patterns reflect how eligibility criteria serve as "common terminology" across protocols – terms like "ECOG 0-1," "measurable disease," "HER2-positive," "EGFR mutation," "MMSE score," "amyloid PET" appear frequently and define the patient cohorts. By leveraging both structured EMR data and contextual interpretation of notes, investigators can match patients to trials effectively, and trial designers use these criteria to balance precision and inclusiveness in enrollment [12] [31].

**References:**


1. Szlezinger K. *et al.* (2023). *Eligibility criteria in breast cancer trials (2020–22 cohort):* upper age limits in 39%, strict comorbidity exclusions in 77%, performance status >1 excluded in 69% [1]. Most trials had lab thresholds for bone marrow (84%), liver (51%), renal (42%) function [3].
2. LUNGevity Working Group (2020). *Lung cancer trial exclusions:* e.g. 90% of trials (2012–2017) excluded patients with creatinine clearance <60 mL/min [8]. Common screen failures were due to brain metastases, prior therapy limits, lab abnormalities, poor performance status [12].

3. Fred Hutch News (2024). *Brain metastases in trials:* by 2022, only ~17% of metastatic cancer trials did not exclude brain metastasis, while ~75% conditionally allowed if treated [6] – showing historic exclusion of CNS disease is slowly easing.
4. Mitchell A. *et al.* (2024). *AD trial criteria analysis:* 87% of trials had age limits, 65% excluded other neurologic disorders, 61% excluded psychiatric disorders [31]. 42% required English fluency [31], reflecting cognitive testing needs. Caregiver involvement is imperative in AD trials [32].
5. FHIR US Core IG & mCODE IG. *Data representation:* Lab results are Observations (grouped by DiagnosticReport) – each Observation = one test and value [22]. DiagnosticReport is used for lab panels, pathology and imaging reports, containing text and references to Observations and images [18] [19]. ECOG performance can be captured as an Observation with standardized coding (LOINC) [10]. These standards enable interoperability of criteria-related data across systems.

---


[1] [2] [3] [25] [26] [28] Eligibility criteria in clinical trials in breast cancer: a cohort study | BMC Medicine | Full Text
https://bmcmedicine.biomedcentral.com/articles/10.1186/s12916-023-02947-y

[4] [5] [7] [8] [9] [11] [12] [13] [14] [15] [16] [21] Expanding Access to Lung Cancer Clinical Trials by Reducing the Use of Restrictiv
Exclusion Criteria: Perspectives of a Multistakeholder Working Group
https://www.lungevity.org/sites/default/files/file-uploads/Expanding%20Access%20to%20Lung%20Cancer%20Clinical%20Trials%20by%20Reducing%20the%20Use%20of%20Restrictive%20Exclusion%20Crit

[6] Clinical trials excluding fewer cancer patients with brain 'mets' | Fred Hutchinson Cancer Center
https://www.fredhutch.org/en/news/center-news/2024/04/clinical-trials-excluding-fewer-cancer-patients-with-brain-metastasis.html

[10] ECOG Performance Status Profile - minimal Common Oncology Data Elements (mCODE) Implementation Guide v4.0.0
https://build.fhir.org/ig/HL7/fhir-mCODE-ig/StructureDefinition-mcode-ecog-performance-status.html

[17] [44] Condition - FHIR v6.0.0-ballot2
https://build.fhir.org/condition-definitions.html

[18] [19] [29] [45] DiagnosticReport - FHIR v6.0.0-ballot2
https://build.fhir.org/diagnosticreport.html

[20] First-in-human phase I dose escalation trial of the first-in-class tumor ...
https://pmc.ncbi.nlm.nih.gov/articles/PMC10787778/

[22] [23] [24] US Core Laboratory Result Observation Profile - US Core Implementation Guide v8.0.0
https://build.fhir.org/ig/HL7/US-Core/StructureDefinition-us-core-observation-lab.html

[27] HER2-Positive Breast Cancer Clinical Trials - Mayo Clinic Research
https://www.mayo.edu/research/clinical-trials/diseases-conditions/her2-positive-breast-cancer/

[30] Modernizing Clinical Trial Eligibility Criteria - ASCO Publications
https://ascopubs.org/doi/10.1200/JCO.2017.74.4102

[31] [37] (PDF) Analysis of eligibility criteria in Alzheimer's and related dementias clinical trials
https://www.researchgate.net/publication/381914955_Analysis_of_eligibility_criteria_in_Alzheimer's_and_related_dementias_clinical_trials

[32] [39] [41] [42] [43] worldwide.com
https://www.worldwide.com/wp-content/uploads/2016/10/Improving-Screen-Failure-and-Recruitment.pdf

[33] [34] [35] [38] TITLE
https://www.siumed.edu/sites/default/files/2023-03/KOHLRUS%20Spring%20Conference%20Presentation%202023.pdf

36 Eligibility Rates among Racially and Ethnically Diverse US ...
https://onlinelibrary.wiley.com/doi/full/10.1002/ana.26819

40 Eligibility rates among racially and ethnically diverse US participants ...
https://pmc.ncbi.nlm.nih.gov/articles/PMC11733864/