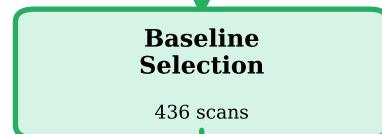# Data Preprocessing & Cleaning Pipeline

Leakage Prevention Steps:

☐ Subject-level de-duplication
☐ Baseline-only selection
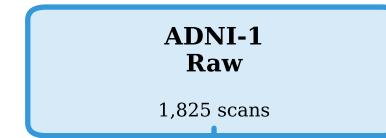☐ Subject-wise train/test splits

**OASIS-1
Raw**

436 scans

**Baseline
Selection**

436 scans

☐ Already
baseline

**CDR 0/0.5
Filter**

205 subjects

-52.8%
(early stage)

**ADNI-1
Raw**

1,825 scans

**Subject
De-duplication**

629 subjects

-65.5%
(unique only)

**Baseline
Selection**

629 subjects

☐ Baseline
only

MRI Preprocessing

Skull strip → MNI152 registration → Normalization

Feature Extraction

ResNet18 2.5D → 512-dim MRI features

OASIS Final

N=205, 512d MRI

ADNI Final

N=629, 512d MRI