Scenario A: fMRI + Behavioral Task

- **Study:** Cognitive control task, 30 participants, fMRI scanning.
- **Pipeline:** Data preprocessed manually in SPM, custom MATLAB scripts for stats.
- **Issues observed:**
    - Motion correction parameters chosen ad hoc.
    - No version control for scripts.
    - Data stored on personal laptop.

Questions for discussion:

- How would you standardize preprocessing?
- What tools/databases could help ensure reproducibility?
- How should data and code be shared?

Scenario B: EEG Study on Sleep

- **Study:** Overnight EEG on 15 participants, investigating sleep spindles.
- **Pipeline:** In-house scripts written over 3 years by different lab members.
- **Issues observed:**
    - Different preprocessing pipelines used across participants.
    - Metadata (electrode placement, sleep stage scoring) missing for some datasets.
    - Results stored as Excel files without raw data links.

Questions for discussion:

- How can the group enforce consistent preprocessing?
- How should metadata be captured and structured?
- What reproducible file formats or standards could be applied?

Scenario C: Multimodal fMRI + PET Study

- **Study:** Combined fMRI + PET to investigate dopamine signaling.
- **Pipeline:** Two separate teams handling each modality, then merging results in R.
- **Issues observed:**
    - No harmonized data structure across modalities.
    - PET preprocessing undocumented.
    - Statistical thresholds chosen post hoc.

Questions for discussion:

- How could the teams coordinate their workflows?
- What reproducibility risks are most serious here?
- Which existing platforms or standards could support multimodal integration?

Step / Pipeline Area	Potential Reproducibility Risks	Possible Solutions / Tools
Data acquisition	e.g. inconsistent scanner parameters, undocumented behavioral tasks	e.g. preregister protocols, standardized acquisition checklists
Preprocessing	e.g. manual corrections, version drift in software	e.g. containerized pipelines (NiPreps, MNE), BIDS-compliant workflows
Analysis	e.g. custom unshared scripts, flexible thresholds	e.g. open-source libraries (Nilearn, SPM), preregistered analysis plans
Data management / sharing	e.g. stored on personal laptops, missing metadata	e.g. OpenNeuro, DataLad, GitHub/Zenodo for code/data