

# Data Management and Reproducibility Standards

Abdullah Lab, IMMEI, University Hospital Bonn

2025-12-20

**Protocol ID:** GEN-DATA-STD-001

**Version:** v1

**Author:** Dillon Corvino

---

## Purpose

This document defines the **mandatory standards** for data management and reproducibility in the Abdullah Lab.

These standards ensure that: - Experiments can be understood and reproduced by others - Data remain interpretable over time - Projects can be transferred without loss of context - Reagents and materials are traceable

All lab members are expected to follow these standards.

---

## Guiding principles

- Clarity over convenience
  - Consistency across the lab
  - One experiment → one clear identifier
  - Raw data are never altered
  - Physical labels must match digital records
- 

## Experiment identifiers

Every experiment must have a **unique experiment ID**.

## Approved formats

Experiment IDs must follow **one of the following formats**:

`<initials><YY>_<NN>`  
`<initials><NN>`

Where: - `<initials>` = experimenter initials (e.g. DC) - `<YY>` = two-digit year (e.g. 25) - `<NN>` = sequential experiment number

## Examples

DC25\_01  
DC25\_02  
DC01  
DC02

The experiment ID must be used consistently across: - Folder names - File names - Lab notebook entries  
- Analysis scripts

---

## File and folder naming

### General rules

- Lowercase letters only
- Underscores (`_`) only
- No spaces or special characters
- Avoid vague terms such as:
  - `final`
  - `new`
  - `test`

### Recommended folder structure

```
experiment_id/  
  raw/  
  processed/  
  analysis/  
  figures/  
  notes/
```

## File naming

Files should **almost always include the experiment ID** to ensure traceability outside the original folder context.

The experiment ID should appear at the **start of the filename**, followed by a short descriptive suffix.

### Examples

DC25\_01\_gating\_strategy.pdf  
DC25\_01\_analysis.wsp  
DC25\_01\_cell\_counts.xlsx  
DC25\_01\_analysis\_notes.md

Files that do not include the experiment ID should be the exception and must still be unambiguous in context.

---

## Approved data storage locations

### Sciebo (approved primary storage)

**Sciebo is the approved storage location for lab data.**

Use Sciebo for: - Raw experimental data - Processed data - Analysis outputs - Shared documents

### Local machines

- Allowed **temporarily only**
- Raw data must be transferred to Sciebo promptly
- Local machines must not be the sole copy of data

### Prohibited storage

- Personal cloud services (e.g. Google Drive, Dropbox)
  - USB sticks as primary or long-term storage
- 

## Raw vs processed data

### Raw data

- Must remain unmodified
- Must retain original file formats
- Must be stored in a dedicated **raw/** directory

## Processed data

- Stored separately from raw data
  - Processing steps must be documented
  - Script-based processing is preferred
- 

## Reagent labelling standards

All reagents must be clearly labelled at the time of preparation or opening.

### Required information on reagent labels

Each reagent container must include: - Reagent name **or** initials / identifier - Sterile or non-sterile status  
- Date opened or prepared - Initials of the person who prepared/opened it

Unlabelled or ambiguously labelled reagents must not be used.

---

## Viral stocks

Viral stocks must be labelled with sufficient information to ensure traceability.

### Recommended format

virus\_strain\_passage\_YYYY-MM-DD

### Example

1cmv\_we\_p3\_2025-02-10

Records must include: - Source - Passage number - Storage location

---

## Cell lines

Cell lines must be labelled in a **similar manner to viral stocks**, ensuring traceability.

## Required information

Cell line labels must include: - Cell line name - Passage number - Date - Initials

### Example

jurkat\_p18\_2025-03-04\_DC

A master record should document: - Cell line source - Genetic modifications (if any) - Thaw date and passage history

---

## Tissue culture flask labelling

All tissue culture flasks must be clearly labelled.

### Required information on flasks

- Cell line name
- Passage number
- Date
- Media used
- Initials

### Example

jurkat | p18 | 2025-03-04 | dmem+penstrep+neaas | DC

---

## Media and reagent modifications

Any modification to base media must be recorded.

### Requirements

- Record all additives (e.g. antibiotics, supplements)
- Use consistent shorthand
- Ensure the formulation is documented digitally

## Example

DMEM + Pen/Strep + NEAAs

If media composition differs from standard lab formulations, this must be clearly documented.

---

## Record keeping

At minimum, records for each experiment must include: - Experiment ID - Date - Operator - Biological material - Reagents used - Deviations from protocol

Electronic or paper lab notebooks are acceptable, but records must be complete and legible.

---

## Versioning

- Use sequential versions: v1, v2, v3
  - Major changes require a new version
  - Deprecated data or reagents must not be deleted without discussion
- 

## Responsibilities

- Each experimenter is responsible for compliance
  - Supervisors may review organisation and records
  - Issues should be corrected promptly when identified
- 

## Notes

These standards are living guidelines and may evolve.

Any ambiguities or gaps should be raised so this document can be updated.