

What to Submit — Contributor Checklist

Cannabis and Health Data Repository

What to Submit

A one-page checklist of what a complete submission to the Cannabis and Health Data Repository looks like. Bring it to your data-management meeting. The Data Coordination Center harmonizes your data to a common data dictionary and manages secure storage and controlled distribution. Your job is to hand off a clean, well-documented, de-identified dataset (we can assist with de-identification if your team needs it).

Required — we cannot begin without these

- ☐ **Your dataset(s).** One or more data files in an accepted format (see below).
- ☐ **A codebook / data dictionary.** Variable names, variable labels, value labels, and how missing values are coded. One codebook that documents every file is acceptable. *(Use the Data Dictionary template in the submission bundle.)*
- ☐ **A study description and documentation.** Enough for another investigator to understand the methodology and data-collection protocol — a study summary, the instruments used, and any user guide or manual.

Strongly encouraged — these materially improve the dataset's long-term value

- ☐ **Programming code** that reproduces derived or constructed variables and links files together.
- ☐ **A blank copy of each data-collection instrument** (survey, CRF, form).
- ☐ **Your IRB approval** and a **blank copy of the consent form(s)** used.
- ☐ **A file inventory** listing everything you are submitting. *(Use the File Inventory template in the submission bundle.)*

Accepted data formats

- ☐ **Preferred** (variable and value labels embedded): **SPSS** · **SAS** (.sas7bdat or .xpt) · **Stata** · **R**
- ☐ **Accepted, with a codebook:** CSV · Excel (.xlsx) · CDISC ODM XML · JSON · Parquet
- ☐ **Talk to us first:** SQLite or other relational-database exports · non-ODM XML
- ☐ **Please export to one of the above first:** JMP · Minitab · MATLAB · Feather · HDF5 · raw SQL dumps

Most platforms already export a preferred or accepted format — REDCap, Castor EDC, OpenClinica, Viedoc, and Medidata Rave export SPSS, SAS, Stata, R, or CDISC ODM XML; Qualtrics, LimeSurvey, QuestionPro, KoboToolbox, and ODK export SPSS, CSV, Excel, or JSON. Export in the format your platform supports best and include the codebook.

Before you send anything

- ☐ You have run the **pre-submission QC checklist** (separate one-pager).
- ☐ You have confirmed **consent permits secondary-research sharing** (see the De-identification one-pager for what your IRB will want).

Questions about whether your data fits? Use the contact page before you begin — we would rather answer a question than receive a submission that has to be reworked.

There is no cost to contribute. Cannabis and Health Data Repository · NIH/NIDA Award 1UM1DA059000 · Johns Hopkins University and Realm of Caring.