

Which Path? — Contribution Decision Aid

Cannabis and Health Data Repository

Which Path?

*A one-page router. There is **one standard contribution path** — a full study contribution that the Data Coordination Center de-identifies, harmonizes, and includes under governed release. Use this to check fit before you begin.*

Step 1 — Is this a fit?

Is your dataset about cannabis or cannabinoid use *and* a health outcome, from human subjects?

- **No** → It may be a better fit elsewhere. Purely genomic or sequence data is usually better served by **dbGaP**; imaging collections by an archive such as **TCIA**. Datasets with no cannabis or health content, and non-human-subjects data, are out of scope. Unsure? Ask us — eligibility can be settled in a brief consultation.
- **Yes** → Continue.

Step 2 — Can it be de-identified?

Can the dataset be de-identified without destroying its analytic value?

- **Unsure / partly** → Contact us. Narrative free text and high-risk indirect identifiers can be routed to the **restricted tier** rather than stripped. Do not silently delete analytically important fields.
- **Yes** → Continue.

Step 3 — Does consent permit sharing?

Does your participant consent permit sharing the data for secondary research, and describe how confidentiality is maintained?

- **Unsure** → Contact us before submitting. Ambiguous consent language is far easier to resolve up front than after data has been transferred.
- **Yes** → You are on the contribution path. Continue to Step 4.

Step 4 — Execute the Data Contribution Agreement

Your institution executes the **Data Contribution Agreement** (*working name; final title pending legal*). Identify your institution's **authorized signing official ahead of time** — the person authorized to sign data agreements is often not the PI, and inviting the wrong signer is the most common cause of a stalled agreement.

Then: the standard contribution path

1. **Apply** — submit a brief description of the candidate dataset and your institution.
2. **DCC review and approval** — the Center confirms scope, de-identifiability, consent, and documentation (distinct from the Steering Committee's review that governs *release* to researchers).
3. **Execute the Data Contribution Agreement** — see Step 4.
4. **Prepare and transfer** — clean and document your data, complete the one institution-side de-identification step, and transfer securely (out-of-band; the web application never receives raw data).
5. **De-identification, harmonization, and your review** — the Center prepares a de-identified, harmonized dataset; you review a preview before it proceeds.
6. **Governed release** — a two-person governance control gates release; the dataset is then accessioned, citable, and available to vetted researchers under an executed Data Use Agreement.

No cost to contribute. Typical preparation and review takes a few months. Cannabis and Health Data Repository · NIH/NIDA Award 1UM1DA059000 · Johns Hopkins University and Realm of Caring.