National Institute of Diabetes and Digestive and Kidney Diseases

Demystifying Data Sharing Plans

James Welch, MGC, CGC

Genetic Counselor

NIDDK Office of the Clinical Director

National Institute of Diabetes and Digestive and Kidney Diseases

Objectives: Understand and Improve Plans

- Understanding what data sharing plans are and are not
- Understanding what data sharing plans should contain
- Understanding NIDDK QI of data sharing plans during quad review

Disclaimer: I focus on helping Intramural Clinical Investigators

Data Management and Sharing Plans

- These are **not** Data Safety and Monitoring Plans
  - DSMP vs. DMSP
  - “Data Sharing Plan” seems more straightforward

National Institute of Diabetes and Digestive and Kidney Diseases

Data Management and Sharing Plans

- Associated with a Z-number (ZIA)
  - Funding and annual reports
    - Now asking you to confirm compliance every year
  - Managed in NIDB
    - NIH Intramural Database
      - data.nidb.nih.gov
  - Controlled by owner of ZIA
    - Can assign a surrogate for the plan/protocol in NIDB
      - E.g., staff clinician is PI or a clinical fellow is taking the lead on writing/submitting the study

What content to focus on?

- Two sections giving the most stress
  1. CDEs/data standards section
  2. Which repository to choose

- Repository:
  - The key for clinical data is <u>controlled-access</u>
    - For researchers, not open to the public
    - Depends on your consent form – always go back to this to confirm compatibility
      - What types of research are allowed on these data?
  - "controlled-access repository **such as _____**"

National Institute of Diabetes and Digestive and Kidney Diseases

Common Data Elements (CDEs)

- Ongoing Protocol vs. New Protocol
  - Ongoing is an option to check
    - Assumption is you're already collecting data
    - Could still note that you'll convert data to applicable NIH-endorsed CDE formats at the time of data upload for sharing, if feasible

- For new protocols this is an opportunity to set yourself up for future data compatibility
  - Between your studies
  - With outside studies that us these same standards, for comparing/validating your data
  - Be ahead of the curve if/when journals start requiring CDEs

DEPARTMENT OF HEALTH & HUMAN SERVICES · USA
NIH National Institute of Diabetes and Digestive and Kidney Diseases

Common Data Elements (CDEs)

- Rather than listing CDE formats as some examples show,
  - Subject ID: Aguk0xu_dB
  - Full Name: EAk1aU4yK
  - Date of Birth: wgPM2Ge1Fq

Consider noting whether you'll be using applicable NIH-endorsed CDEs

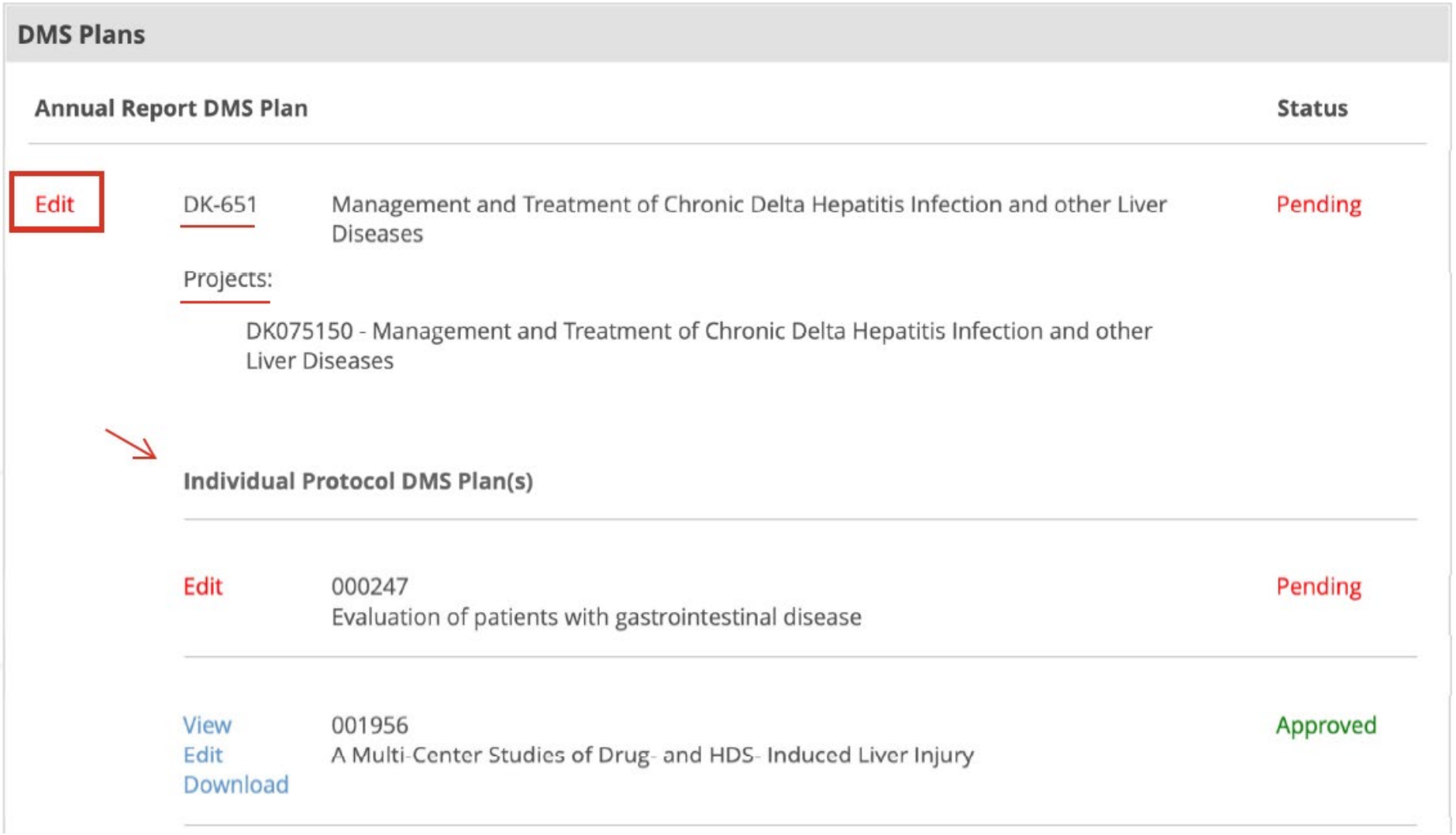
- If yes: focus on this when building your RedCap project, study surveys, or other data collection instruments
- If no: note that it's an ongoing protocol or part of a series of studies for which backwards compatibility of data is the priority

National Institute of Diabetes and Digestive and Kidney Diseases

NIDDK IRP DMSP Quality Improvement

- At Quadrennial Review of the protocol
- Main Goals:
  - Have a separate DMSP for each protocol
  - … that makes sense for that study based on the areas of focus we discussed
- In the past, there was often a single DMSP for the entire ZIA

Individual Protocol DMSPs



Individual Protocol DMSPs

Add another protocol:

Enter Protocol Number

Enter Protocol Title (Required)

Info on Standards

Describe what Common Data Elements (CDEs) will be used. Justify if CDEs are not used.

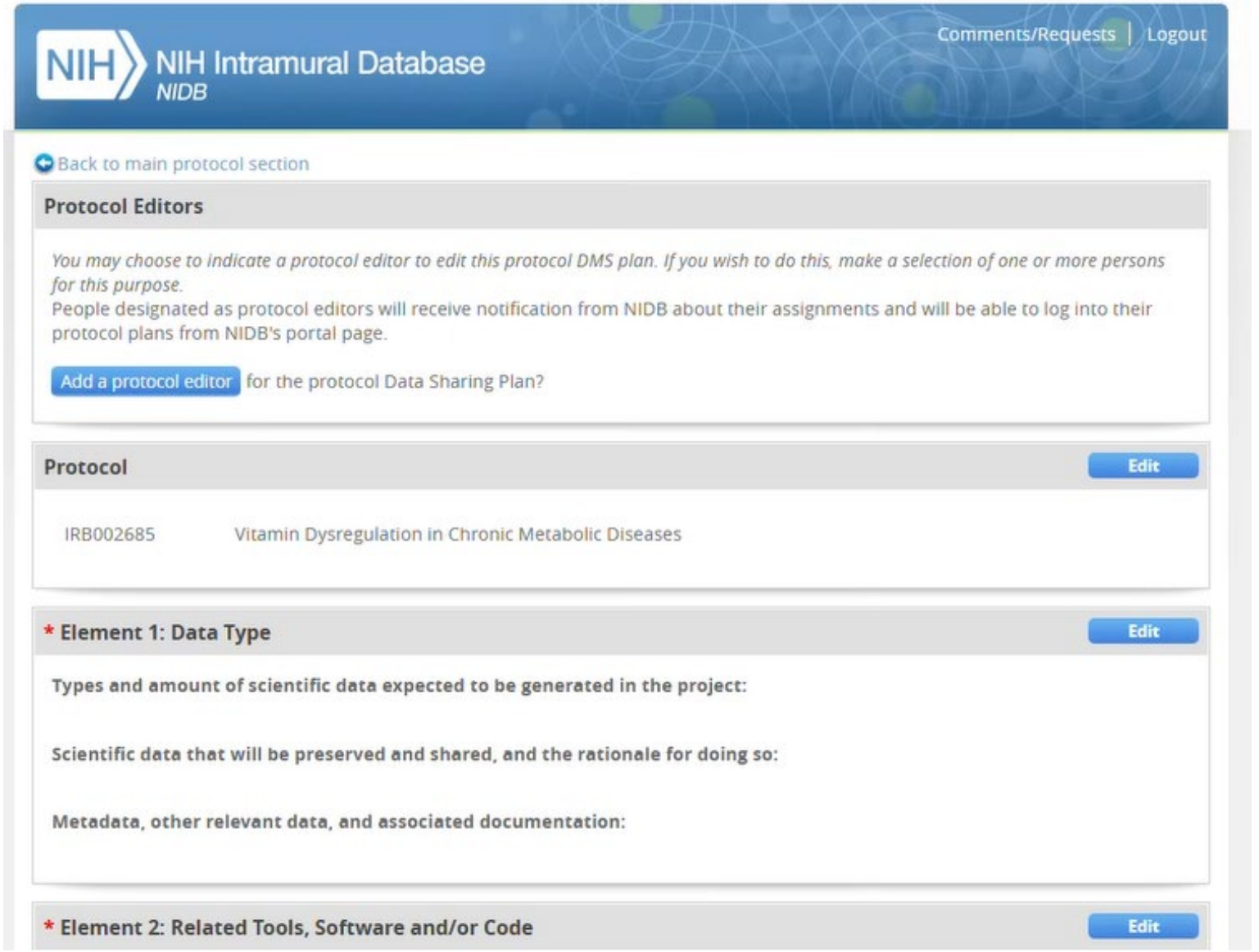
CDEs allow data to be collected in the same way across multiple research studies and are defined unambiguously in both human and machine-computable terms. NIH has established a "Common Data Element (CDE) Resource Portal" (http://cde.nih.gov/) to help investigators identify CDEs when collecting data, particularly for clinical research, and when developing protocols and case report forms.

☐ This is an approved/ongoing protocol.
(Checking this box allows you to leave this text box blank. Otherwise, entry is required for the protocol's CDE.)

☐ Check here if you would like to create a DMS plan specific to this protocol. If not checked, you should combine the DMS plan for this protocol with the DMS plan for the ZIA as a whole.

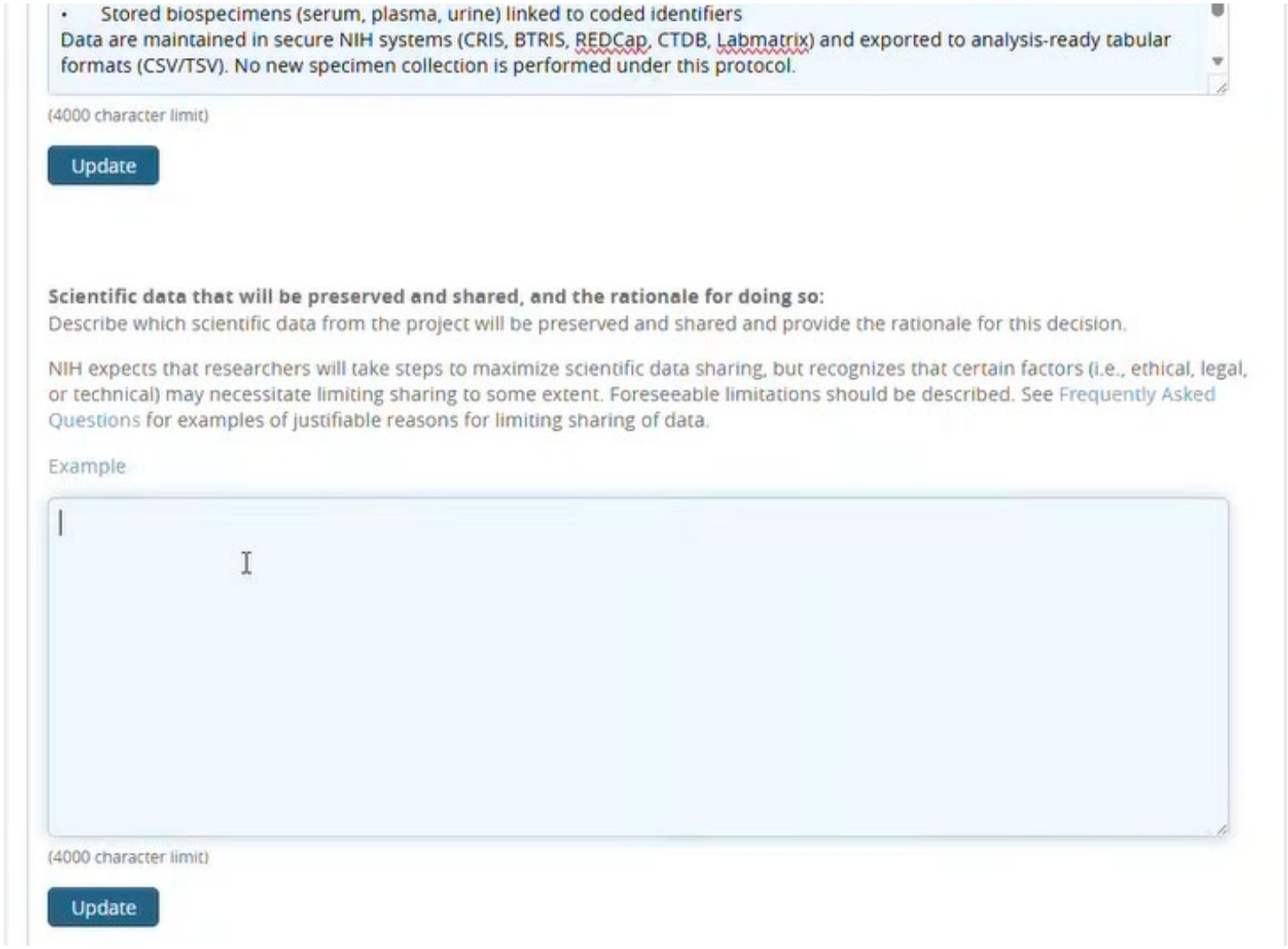
Save and Continue

Tips and Tricks: Assigning a surrogate “Protocol Editor”



Tips and Tricks: Multi-part Elements

For multi-part elements: paste your answer to one part and select Update before pasting the next part – you can only update one part at a time



Tips and Tricks: Element 3 Formatting

Element 3: Data Standards:

Data standards refer to community-accepted approaches to methods of organizing, documenting, and formatting data to aid in data aggregation, sharing and reuse, and could refer to common data elements (CDEs), data dictionaries, data models, vocabulary, etc. Metadata standards (e.g., DataCite, DublinCore) are also data standards. Determining which standards are relevant to specific data may depend on the data repository where the data will be submitted as well as the standards adopted by the specific scientific community and/or by specific NIH ICs. Refer to: https://fairsharing.org/ that makes over 1600 standards searchable.

Describe what Common Data Elements (CDEs) will be used. Justify if CDEs are not used.

CDEs allow data to be collected in the same way across multiple research studies and are defined unambiguously in both human and machine-computable terms. NIH has established a "Common Data Element (CDE) Resource Portal" (http://cde.nih.gov/) to help investigators identify CDEs when collecting data, particularly for clinical research, and when developing protocols and case report forms.

☐ This is an approved/ongoing protocol.
(Checking this box allows you to leave this text box blank. Otherwise, entry is required for the protocol's CDE.)

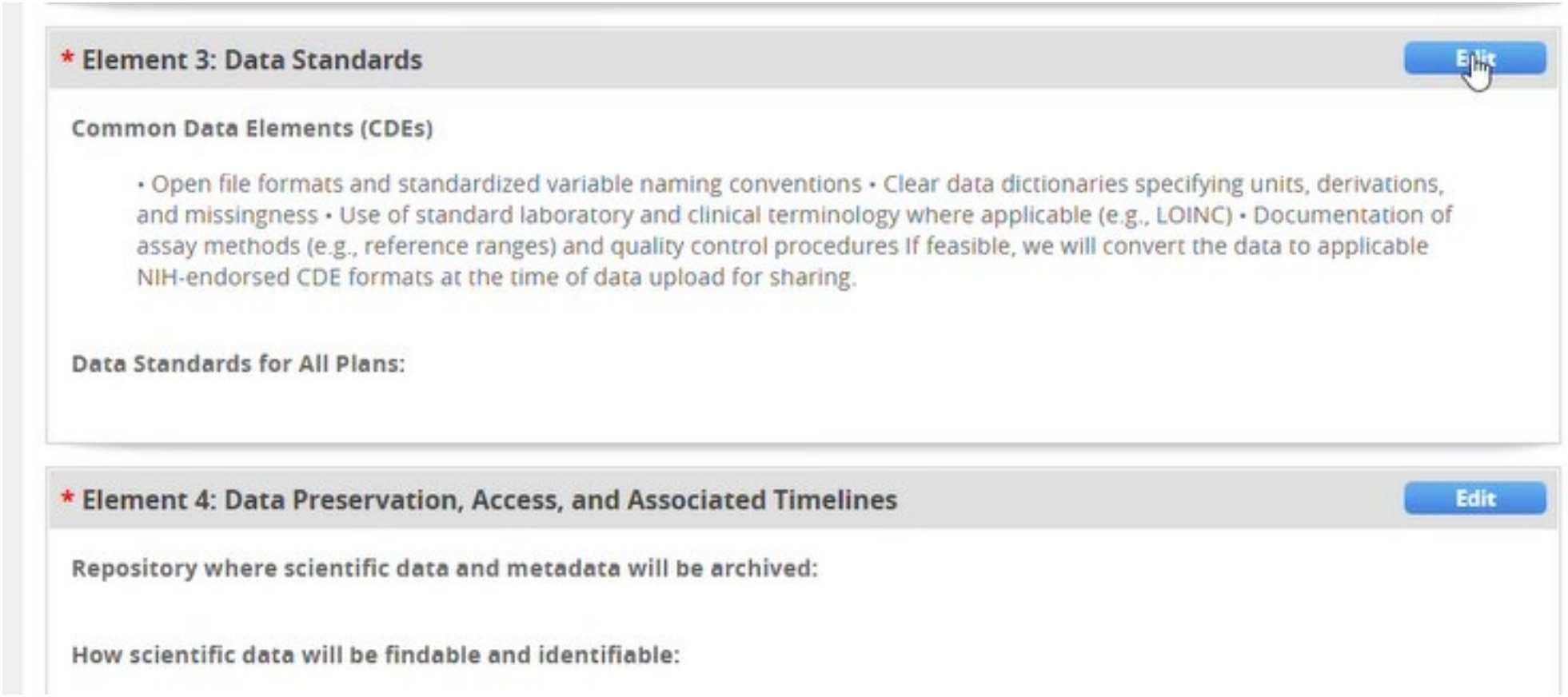
- Open file formats and standardized variable naming conventions
- Clear data dictionaries specifying units, derivations, and missingness
- Use of standard laboratory and clinical terminology where applicable (e.g., LOINC)
- Documentation of assay methods (e.g., reference ranges) and quality control procedures

If feasible, we will convert the data to applicable NIH-endorsed CDE formats at the time of data upload for sharing.

(4000 character limit)

Update

Tips and Tricks: Element 3 Formatting



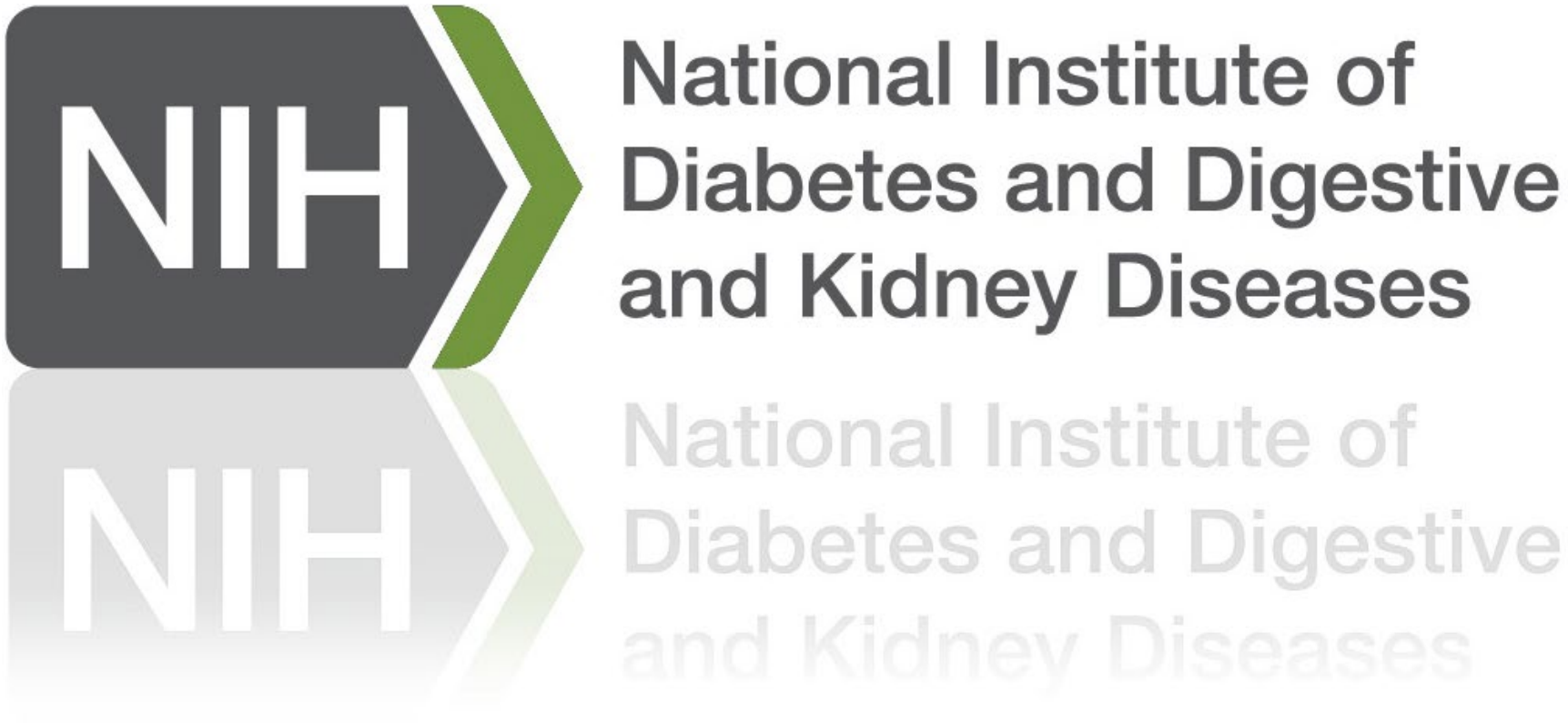
- This section doesn't handle formatting well (i.e., bullet points), but all other text boxes do, including part 2 of Element 3 if you want to list your CDEs there instead.

Genomic Data

- If you're an NIDDK IRP clinical investigator, contact me to help with
  - Protocol/consent language for genetics/exome/genome sequencing
  - Addressing genomic data in your data sharing plan
    - Different repositories for different types of data
    - Make sure your study qualifies as generating large genomic data sets

james.welch@nih.gov

National Institute of Diabetes and Digestive and Kidney Diseases


NIH
National Institute of
Diabetes and Digestive
and Kidney Diseases