

Data or words?

What USDM holds for objectives, endpoints, eligibility criteria and the schedule

Data

Data in a string

Words

The question

For each design element: which parts are data, and which are still words?

AN OBJECTIVE IN USDM · CDISC PILOT STUDY EXAMPLE

Level: secondary objective

Its 3 endpoints, linked

“To assess the dose-dependent improvement in behavior. Improved scores on the Revised Neuropsychiatric Inventory (NPI-X) will indicate improvement in these areas.”

What is measured, and how it is judged, sits inside the sentence.

Data

Choices from a list, numbers, links. A system can check and reuse them.

Data in a string

An identifier or a duration written as text. Usable only if you know the convention.

Words

Sentences and written rules. A person has to read them.

Each reading here comes from USDM v4 represented as a knowledge graph.

Five design elements: the schedule is data, the rest is mostly words

Data Data in a string Words

Objective ↗

Level Its endpoints

The objective, written



Endpoint ↗

Level

The endpoint, written Its purpose

no link
in USDM

Eligibility criterion ↗

Inclusion or exclusion Order

Label like "I1"

The criterion, written

no link in USDM

Protocol document, its sections ↗

Schedule of activities ↗

	Visit	Visit	Visit	Visit	visit type, contact mode
Activity	•	•		•	what is measured
Activity		•	•	•	timing and windows
Activity	•			•	Conditions for branching
					Visit start and end rules

Arms × epochs ↗

	Screening	Treatment	Follow-up
Arm 1	element	element	element
Arm 2	element	element	element

Inside each element: what happens, start rule, end rule

In short

USDM gives every design element an identity. Much of the content is still words.

Identity, for the objective on slide 2

USDM, the kind name Objective · C-code C142450 (Study Objective)

USDM, this one id Objective_3 · name OBJ3 · unique within the study file

usdm-rdf adds a global identifier for the kind, linked to NCI C142450: <https://w3id.org/cdisc/usdm/v4/Objective>

Held as data

The schedule of activities, and the arm and epoch grid

Data in a string

Timing values and windows; labels like “I1”

Held as words

Objectives, endpoints, eligibility criteria, rules inside elements

Not held

Endpoint to assessments; design element to protocol section

Beyond USDM

The words are not a dead end

A knowledge graph can represent the entities and relationships from study design to data capture. Structuring what the words say in such a graph turns them into data, and that is where the link from an endpoint to its assessments can live.