

Data Dictionaries

1. Costs per procedure

This is the schema of the costs for individual procedures e.g. “urinalysis”, used to build the budget from the schedule of events

Field Name	Data Type	Description	Example
id	Integer	Unique primary identifier for the reference cost record.	6658
procedure	String	Full text description or official title of the medical service or procedure.	Beratung – auch mittels Fernsprecher
cost	Decimal / Float	The monetary cost associated with the medical procedure (e.g., in EUR).	10.72
created_at	Timestamp	Record creation date and time (YYYY-MM-DD HH:MM:SS.SSS).	2025-10-17 07:34:53.090
updated_at	Timestamp	Record last modification date and time (YYYY-MM-DD HH:MM:SS.SSS).	2025-10-17 07:34:53.090
metadata	JSON / Object	Additional source references including URL	{"url": "https://...", "source": "GOÄ Ziffern..."}

2. Top-Down Cost Data Dictionary

This schema describes the dataset structure for the dataset of entire clinical trial costs.

Field Name	Data Type	Description	Example
nct_id	String	ClinicalTrials.gov unique study identifier.	NCT04167540
cost	Integer	Total estimated trial cost	546046200
cost_source_url	String (URL)	Direct URL to the source	https://www.cirm.ca.gov/...
other_data	Object	Nested metadata containing other custom study details	<code>{"name_cirm": "...", "phase_cirm": "Phase 1"}</code>
phase	Integer	Standardized numeric trial phase identifier.	1
title	String	Short public title of the clinical study.	GDNF Gene Therapy for Parkinson's Disease
description	String	Detailed summary of trial objectives and methodology.	The objective of this Phase 1b investigation...
conditions	Array of Strings	Target medical conditions or diseases evaluated.	<code>["Parkinson's Disease", "Parkinson Disease"]</code>
interventions	Array of Strings	Interventional drugs, therapies, or biological products tested.	<code>["AAV2-GDNF"]</code>
conditions_joined	String	Comma-separated concatenation of targeted conditions.	Parkinson's Disease, Parkinson Disease
interventions_joined	String	Comma-separated concatenation of tested interventions.	AAV2-GDNF
intervention_type	String	High-level categorization of the	drug

Field Name	Data Type	Description	Example
		intervention type.	
condition_codes	Array of Arrays	Medical Subject Headings (MeSH) mapping tuples: [MeshObject, StartPos, EndPos].	[[{"mesh_id": "D016631", ... }, 4, 5]]
ctgov_url	String (URL)	Direct link to the public record on ClinicalTrials.gov.	https://clinicaltrials.gov/study/NCT04167540
enrollment	Integer	Target or actual total participant count.	11
country_codes	Array of Strings	ISO Alpha-2 country codes where the study takes place.	["US"]
lmic	Boolean	Flag indicating if the trial is in a Low- and Middle-Income Country.	false