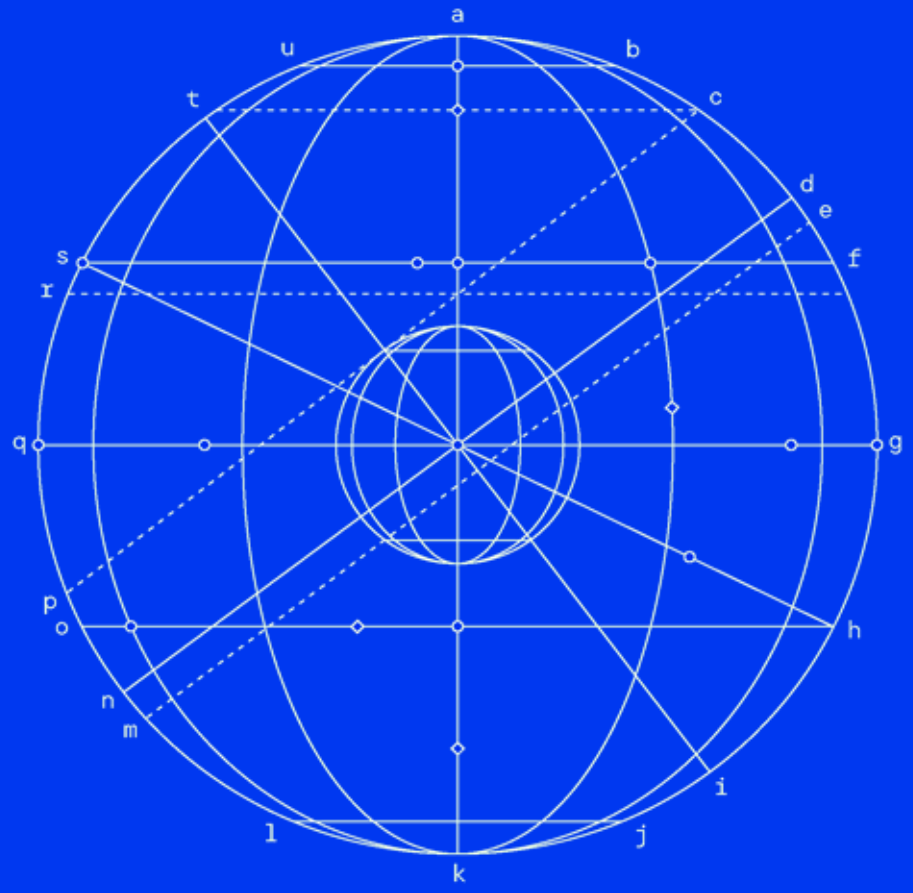
Operationalizing CDISC Biomedical Concepts as the Semantic Layer of an eSource Platform

EHR-to-EDC data transfer, AI-grounded extraction and metadata-driven study configuration from one concept library

PHUSE Real World Data Autumn Event 2026

Paradigm Health



Three problems had to be solved to enable eSource at scale

Paradigm sought to build its eSource capability from the ground up with biomedical concepts at the core – solving these problems were key to making this possible

PROBLEM 01 Semantics

CRFs are not machine-readable

Fields are string identifiers. Nothing tells a system that SYSBP on one form and VS_SYS_1 on another are the same measurement, or what unit it expects.

WHAT PARADIGM BUILT

Biomedical Concepts as the semantic layer: every field is bound to a BC and a Data Element Concept, so the pipeline knows where EHR data goes and how to normalize it.

PROBLEM 02 Structure

Sponsor CRFs arrive as they are

Repeating item groups, single groups, tables, free-form logs. Sponsors will not redesign their casebook to fit our data model, so the model has to fit theirs.

WHAT PARADIGM BUILT

A study builder that applies BCs and DECs to any form structure, and a digital protocol that tells the pipeline which instance belongs in which visit.

PROBLEM 03 Scale

Perfect definitions do not scale

Waiting for a complete, agreed definition of every activity and concept is a bottleneck. Studies start before the standards catch up.

WHAT PARADIGM BUILT

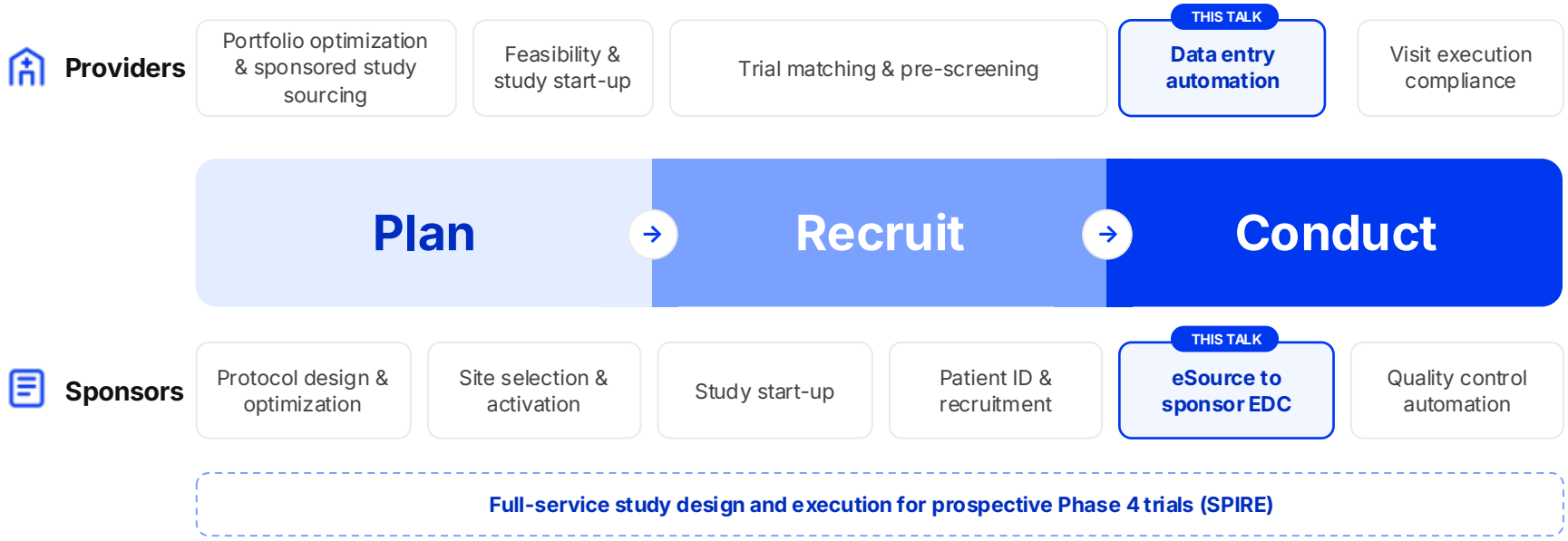
Composable modeling: canonical domains and DECs compose into BC classes, BCs are printed from classes, and activities are supersets rather than exact lists.

© 2026 Paradigm Health, Inc.

Paradigm Health

Paradigm Health connects providers and sponsors across the trial lifecycle

This talk lives in the Conduct phase: eSource auto-population of case report forms from EHR data



© 2026 Paradigm Health, Inc.

Paradigm Health

PROBLEM 01 · SEMANTICS

Case report forms are not machine-readable

What Paradigm built: Biomedical Concepts as the semantic layer, so every field carries meaning a pipeline can act on.

© 2026 Paradigm Health, Inc.

Within an EDC, a CRF field is just a string

Field identifiers carry no semantics, so nothing tells a pipeline which concept a field captures, where an EHR observation belongs, or how to normalize it

TODAY: WHAT THE SYSTEM SEES ON A VITAL SIGNS FORM

FIELD 01	SYSBP	Enter value
FIELD 02	VS_DIA	Enter value
FIELD 03	**Pulse (bpm)**	Enter value

???

Three unrelated strings

No unit, no data type, no link to a source observation, no relationship to each other. Every mapping is a human decision, repeated per study.

WITH BIOMEDICAL CONCEPTS: WHAT THE SYSTEM NEEDS

Parent BC · Vital Signs — links the three concepts below

FIELD 01	SYSBP	Enter value	BC Systolic Blood Pressure · C25298 DEC Result Value · mmHg
FIELD 02	VS_DIA	Enter value	BC Diastolic Blood Pressure · C25299 DEC Result Value · mmHg
FIELD 03	**Pulse (bpm)**	Enter value	BC Pulse Rate · C49677 DEC Result Value · beats/min

3

Three concepts, one parent

Each field resolves to its own Biomedical Concept, and all three link to the Vital Signs parent BC. A FHIR Observation coded LOINC 8480-6, 8462-4 or 8867-4 now finds its field, its unit and its code list.

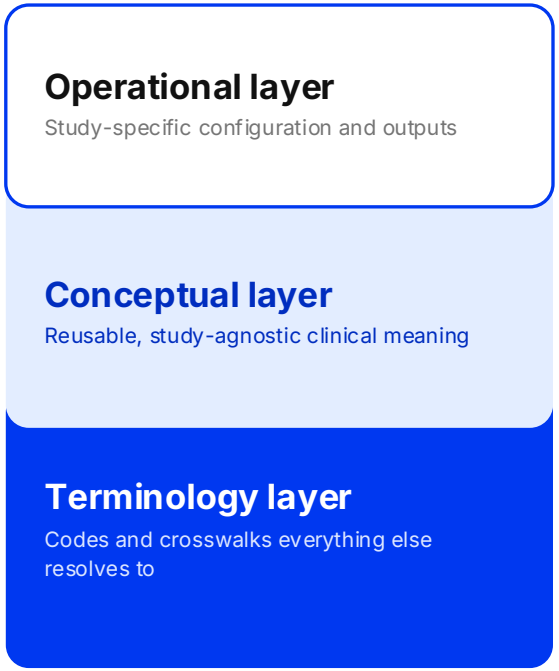
Coding the form is what lets structured FHIR data find its way onto the right form, the right field, in the right unit.

© 2026 Paradigm Health, Inc.

Paradigm Health

Three metadata layers make a form machine-readable

Each layer is the foundation for the next; Paradigm required this architecture to make BC management scalable



OPERATIONAL

`eCRF configuration` `Casebook versions` `SDTM` `ADaM` `define-xml`

Printed from the layers below. Rebuilt per study, never hand-mapped.

CONCEPTUAL

`Activity Concepts` `Biomedical Concepts` `BC Class templates` `Canonical Domains` `Data Element Concepts`

Where this talk lives. A knowledge graph of what should be observed and how the pieces relate.

TERMINOLOGY

`Base terms` `Crosswalks` `Value sets`

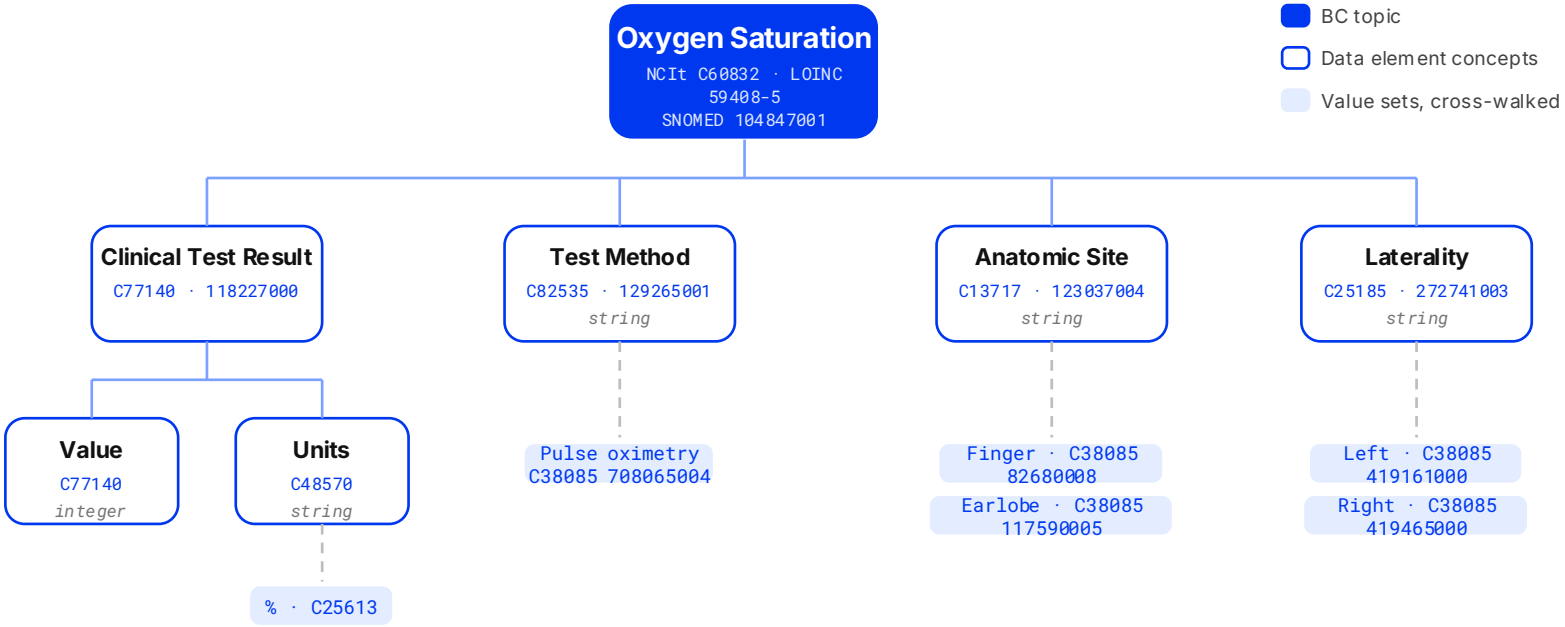
NCIt, LOINC, SNOMED CT, UCUM, RxNorm and sponsor code lists, linked so any code can be resolved to any other.

© 2026 Paradigm Health, Inc.

Paradigm Health

A Biomedical Concept is a unit of knowledge: what should be observed, and how

Oxygen Saturation as a graph of data element concepts, each coded and cross-walked, with value sets at the leaves

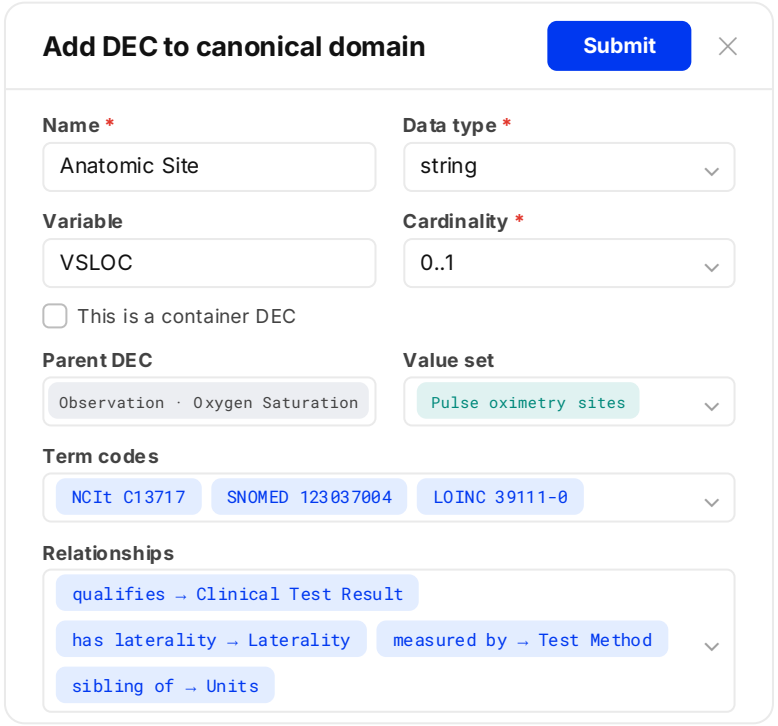


© 2026 Paradigm Health, Inc.

Paradigm Health

The anatomy of a Data Element Concept

Every attribute is something the pipeline or the study builder acts on



Name & term codes	**Meaning the graph can resolve.** The codes are the join key to FHIR, CDASH and OMOP through crosswalks, so an EHR value finds this element by concept, not by label.
Data type	**Typed extraction and validation.** Integers, strings, dates and codeable concepts are parsed differently from FHIR and validated before they land on a form.
Variable	**A path to SDTM.** The CDASH or SDTM variable this element feeds, so submission datasets are derived rather than mapped by hand.
Cardinality	**Single value or repeating.** Tells the study builder whether an element can populate a log or a table row.
Container & parent	**Nesting.** A result is a container for value and unit; the hierarchy is what makes the BC a graph rather than a flat list.
Relationships	**Edges in the knowledge graph.** What this element qualifies and what qualifies it, so a pipeline can traverse from a result to its site, method and laterality.
Value set	**Controlled terms with crosswalks.** Allowed answers, each coded, so sponsor code lists can be normalized to and from them.

© 2026 Paradigm Health, Inc.

Paradigm Health

PROBLEM 02 · STRUCTURE

Sponsor case report forms arrive as they are

What we built: a study builder that applies BCs and DECs to any form structure, and a digital protocol that tells the pipeline which instance belongs in which visit.

© 2026 Paradigm Health, Inc.

Study Builder maps concepts onto any form structure

Single item groups, repeating item groups, tables and mixed layouts, all bound to the same BCs and DECs

SINGLE ITEM GROUP `1 : 1`

Vital Signs

Activity · Vital signs

Systolic BP	118	C25298
Diastolic BP	76	C25299
Pulse	68	C49677
SpO_2	97	C60832

One record per visit

REPEATING ITEM GROUP `1 : 0..*`

Concomitant Meds

Activity · Con meds

#	Medication		Dose	Route
1	Metformin	RxNorm	500 mg	Oral
2	Lisinopril	RxNorm	10 mg	Oral
3	Atorvastatin	RxNorm	20 mg	Oral

+ Add row · one BC instance per row

TABLE LAYOUT `rows × columns`

Hematology

Activity · Hematology

Test		Result	Unit	Date
Hematocrit	C64796	42.3	%	23 Dec
Leukocytes	C51948	7.8	$10^3/\mu L$	23 Dec
Platelets	C51951	245	$10^3/\mu L$	23 Dec

Each row is a BC; each column a DEC

Same concepts, different structures

The pipeline routes data by BC and DEC, not by field position or label, so the sponsor's form stays as the sponsor designed it.

Single group · Repeating group · Table · Mixed

© 2026 Paradigm Health, Inc.

Paradigm Health

The digital protocol plus BCs find the data most likely to be study-specific

Protocol timing narrows the record to a visit window; the BC graph narrows the codes; the best instance pre-populates the eCRF

1 The digital protocol says when

Schedule of Activities

	Epoch	Screening	Cycles 1-2	
	Visit Parent			
	Visit	D1	D1	D8, D15, D22
Activities	Window	±3	±3	±1
Vital Signs		☑	☑	☑
Height		☑	☐	☐
Weight		☑	☑	☐
12-lead ECG		☑	As clinically indicated	
Hepatitis Serology Testing		☑	☐	☐
HBV DNA		☑	If indicated, every 12 weeks	

Schedule of activities: Vital Signs at Cycles 1–2, Day 8, window ±1 day.

2 The BC says what qualifies



Filter Observations coded to the BC topic (LOINC 59408-5 and crosswalks)

Rank Inside the window, then closest to the target date

Protocol timing narrows the record; the BC graph narrows the codes.

3 The eCRF is pre-populated

Vital Signs · C1D8 `Auto-filled`

SpO_2	**97 %** ✓
Method	**Pulse oximetry** ✓
Site	**Finger** ✓
Laterality	**Left** ✓
Date	**23 Dec 2024** ✓

Source-linked to the FHIR observation for review and audit.

Site staff verify rather than transcribe.

© 2026 Paradigm Health, Inc.

Paradigm Health

Pipeline example: hematology from FHIR to eCRF

Ingest, extract BC instances, group into activity instances, import

1 Ingest EMR FHIR data → **2 Extract BC instances** → **3 Group into activity instances** → **4 Site imports into the eCRF**

- **Observation**
- **DiagnosticReport**
- **Specimen**

Pulled per subject, per visit window, from the provider's FHIR API.

BC	Collected	Value	Unit
Hematocrit	23 Dec 10:30	42.3	%
Leukocyte count	23 Dec 10:30	7.8	10^3/µL
Leukocyte count	23 Dec 11:25	6.4	10^3/µL
Platelet count	23 Dec 11:25	245	10^3/µL

Each LOINC-coded observation resolves to its own BC through the crosswalk; units are normalized to the BC's value set.


```
{ "activityConcepts": [
  { "name": "Hematology",
    "instanceDate": "2024-12-23T10:30",
    "biomedicalConceptInstances": [
      { "test": "Hematocrit",
        "value": 42.3, "unit": "%" },
      { "test": "Hemoglobin",
        "value": 13.9, "unit": "g/dL" },
      { "test": "Leukocyte Count",
        "value": 7.8, "unit": "10³/µL" },
      { "test": "Erythrocyte Count",
        "value": 4.6, "unit": "10⁶/µL" } ]
},
  { "name": "Hematology",
    "instanceDate": "2024-12-23T11:25",
    "biomedicalConceptInstances": [
      { "test": "Leukocyte Count",
        "value": 6.4, "unit": "10³/µL" },
      { "test": "Platelet Count",
        "value": 245, "unit": "10³/µL" },
      { "test": "Neutrophil Count",
        "value": 3.9, "unit": "10³/µL" } ]
} ] }
```


Grouped by specimen collection time into activity instances.



Candidate instances side by side; one click imports the chosen one.

© 2026 Paradigm Health, Inc.

Paradigm Health

From codeable concept to sponsor code list

Normalization happens twice: once to the canonical term, once to the sponsor's controlled terminology



```
"code": { "coding": [ {
    "system": "http://loinc.org",
    "code": "26464-8",
    "display": "Leukocytes Bld NCnc Pt Qn" } ],
  "text": "WBC" },
"subject": { "reference": "Patient/PD-1042" },
"effectiveDateTime": "2025-01-07",
"valueQuantity": {
  "value": 7.5,
  "unit": "K/uL",
  "system": "http://unitsofmeasure.org",
  "code": "10*3/uL" },
"referenceRange": [ {
  "low":  { "value": 4.5,  "code": "10*3/uL" },
  "high": { "value": 11.0, "code": "10*3/uL" } }
]
```

EHR CODEABLE CONCEPT

10*3/uL

UCUM code as sent by the provider's FHIR server. Another site may send "K/uL" or "thou/mm3".

→

CANONICAL TERM

Billion per liter

Curated semantic crosswalk to the CDISC / NCIt unit the BC's value set expects.

→

SPONSOR CODE LIST

X10³/µL

Sponsor-specific mapping applied at the form, so the EDC receives its own controlled term.

INFORMATICS-MANAGED NORMALIZATION

Terminology & crosswalks

- Curate canonical terms, value sets and aliases
- Map EHR and local codes to CDISC / NCIt once, reuse everywhere

Sponsor delivery

- Maintain sponsor code-list and unit mappings per study
- Review exceptions, then version and publish reusable maps

Crosswalks are curated once by informatics and reused across every study and sponsor.

© 2026 Paradigm Health, Inc.

Paradigm Health

PROBLEM 03 · SCALE

Perfect definitions do not scale

What we built: composable modeling with supersets and generic concepts, so studies never wait on a finished library.

© 2026 Paradigm Health, Inc.

Paradigm's activity concepts are supersets, not definitions

Define everything a vital-signs activity could collect once; each study enables the subset it needs

Activity concept · Vital Signs

13 BCs

- Systolic blood pressure
- Diastolic blood pressure
- Pulse
- Body mass index
- Frame size
- Heart rate
- Temperature
- Waist circumference
- Respiratory rate
- Oxygen saturation
- Head circumference
- Lean tissue mass
- Total body water

Defined once as everything a vital-signs activity could collect. Adding a BC later never breaks an existing study.

Study A · Oncology

7 enabled

- [x] Systolic BP
- [x] Diastolic BP
- [x] Pulse
- [ ] Body mass index
- [ ] Frame size
- [x] Heart rate
- [x] Temperature
- [ ] Waist circ.
- [x] Respiratory rate
- [x] Oxygen saturation
- [ ] Head circ.
- [ ] Lean tissue mass
- [ ] Total body water

Configured in minutes from the superset. The pipeline only extracts what is enabled.

Study B · Cardiology

6 enabled

- [x] Systolic BP
- [x] Diastolic BP
- [ ] Pulse
- [x] Body mass index
- [ ] Frame size
- [x] Heart rate
- [ ] Temperature
- [x] Waist circ.
- [ ] Respiratory rate
- [ ] Oxygen saturation
- [ ] Head circ.
- [ ] Lean tissue mass
- [x] Total body water

Same activity concept, different subset. No debate about the "correct" definition of Vital Signs.

No study is blocked on agreeing the one correct definition of an activity.

© 2026 Paradigm Health, Inc.

Paradigm Health

Generic BCs: model the shape, not every instance

One Concomitant Medication BC covers every medication the EHR can send

ONE BC PER DRUG?

- BC ~~Metformin use~~
- BC ~~Lisinopril use~~
- BC ~~Atorvastatin use~~
- BC ~~Pembrolizumab use~~
- BC ~~Ondansetron use~~

… × every product in RxNorm

Modeling at the level of the individual medication means the library is never finished, and every new study waits on new definitions.

Generic BC · Concomitant Medication

BC class · Medication use

Medication RxNorm · NCIt *codeable concept*	**Dose** CMDOSE *decimal*	**Dose unit** CMDOSU · UCUM *value set*	**Route** CMROUTE *value set*
Frequency CMDOSFRQ *value set*	**Start date** CMSTDTC *date*	**End date** CMENDTC *date*	**Indication** CMINDC · SNOMED *codeable concept*

Instance	Medication	Dose	Route	Frequency	Start	Source
1	Metformin · RxNorm 6809	500 mg	Oral	BID	2023-04-02	MedicationStatement
2	Lisinopril · RxNorm 29046	10 mg	Oral	QD	2022-11-15	MedicationRequest
3	Ondansetron · RxNorm 7804	8 mg	IV	PRN	2024-12-23	MedicationAdministration

The drug itself is a value of the Medication element, resolved through RxNorm, not a concept to be modeled. One BC covers every medication the EHR can send.

Generic concepts absorb variation the library never has to model explicitly.

© 2026 Paradigm Health, Inc.

Paradigm Health

Scalable BC Creation: compose from a canonical base, then print

Canonical domains and data element concepts compose into BC classes; BCs are printed from classes; activity concepts group them

1 BASE

Canonical domains

Standards-agnostic domains modeled on FHIR resources and implementation guides, triangulated with CDASH and OMOP.

- Observation
- Medication
- Condition
- Specimen

2 ELEMENTS

Data element concepts

Typed, coded attributes that live in a canonical domain and can be reused by any concept in it.

- Result
- Unit
- Method
- Site
- Laterality
- Date

3 TEMPLATES

BC classes

DECs composed into a reusable shape. A class says which elements a family of concepts shares.

- Vital sign measurement
- Lab test result
- Medication use

4 PRINTED

Biomedical concepts

A BC is a class plus a topic code and value sets. Printing a new BC is minutes, not a modeling project.

- Systolic BP
- SpO_2
- Heart rate
- Platelet count

5 GROUPED

Activity concepts

A superset of every BC that could be part of an activity. Each study enables the subset it collects.

- Vital signs
- Hematology
- Concomitant meds

REFERENCE MODELS TRIANGULATED

- FHIR R4
- US Core & IPS guides
- CDASH
- OMOP CDM

TERMINOLOGIES CROSS-WALKED

- NCIt
- LOINC
- SNOMED CT
- UCUM
- RxNorm
- Sponsor code lists

© 2026 Paradigm Health, Inc.

Paradigm Health

ADDITIONAL BENEFITS

BCs create value beyond structured EHR data transfer

© 2026 Paradigm Health, Inc.

The same semantic layer powers more than eSource

Paradigm's eSource Casebook capability can leverage BC-linked eCRFs to drive value beyond data entry automation

AI-grounded extraction

BC value-level metadata becomes the structured-output schema for LLM pipelines over notes and documents. The model can only answer in coded terms the graph already knows, and the BC graph doubles as its semantic memory.

`Structured outputs` `Semantic memory` `Coded, not free text`

Semantically aware forms

Because every field knows its concept, the casebook can render a subject's data as timelines and trends across visits, navigate by concept instead of by page, and flag out-of-range or inconsistent values at entry.

`Cross-visit trends` `Concept navigation` `Entry-time checks`

Standards-driven outputs

Each DEC carries its CDASH and SDTM variable, so datasets, define-xml and visualizations are derived from the same concept library that configured the study. One definition, every downstream use.

`SDTM` `define-xml` `Visualization`

Every capability reads the same library. None needs its own mapping layer.

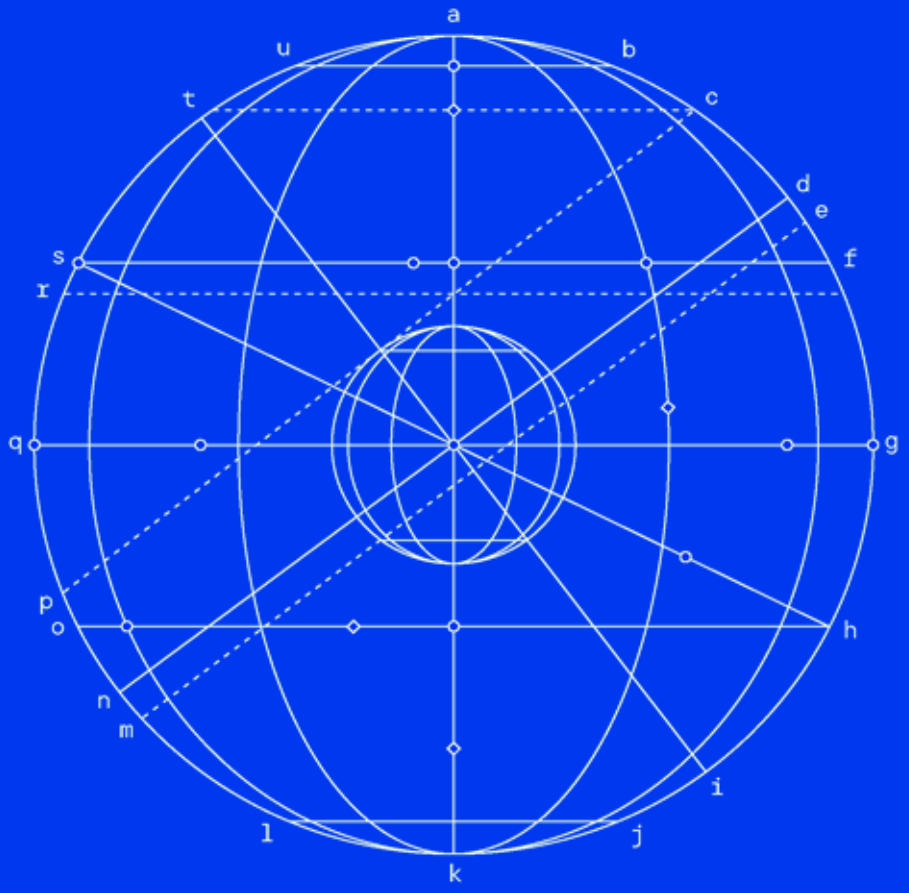
© 2026 Paradigm Health, Inc.

Paradigm Health

Thank you

Questions and discussion

- CRFs made machine-readable with BCs
- Any sponsor form, mapped by concept
- Composable modeling that scales



Paradigm Health