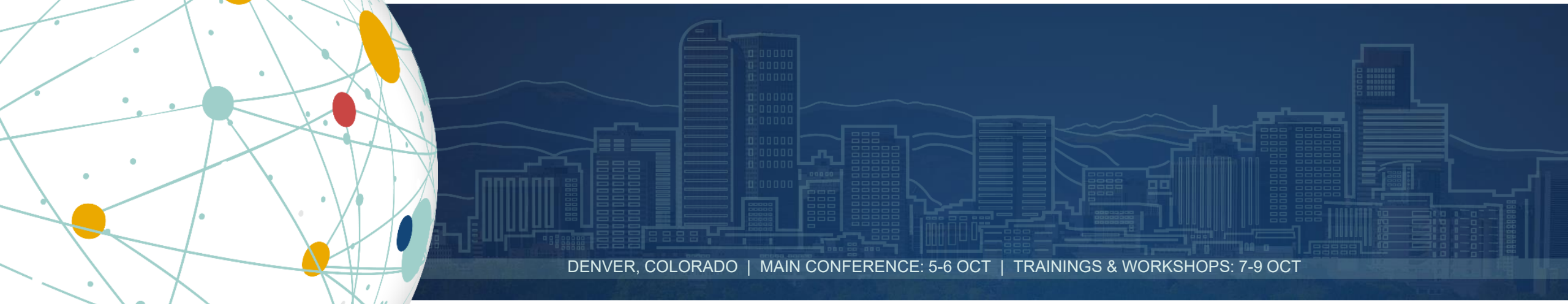


2026 US Interchange

THE FUTURE IS CONNECTED: STANDARDS AND AI POWERING DIGITAL TRANSFORMATION



Semantic Bridges: Linking Study Design to Analytical Concepts

Luke Morgan, Senior Manager, Clinical Data Standardization, Medidata
Bhargav Koduru, Director, Data Curation, Medidata

Speakers



Luke Morgan
Senior Manager, Clinical Data Standardization
Medidata



Bhargav Koduru
Director, Data Curation
Medidata

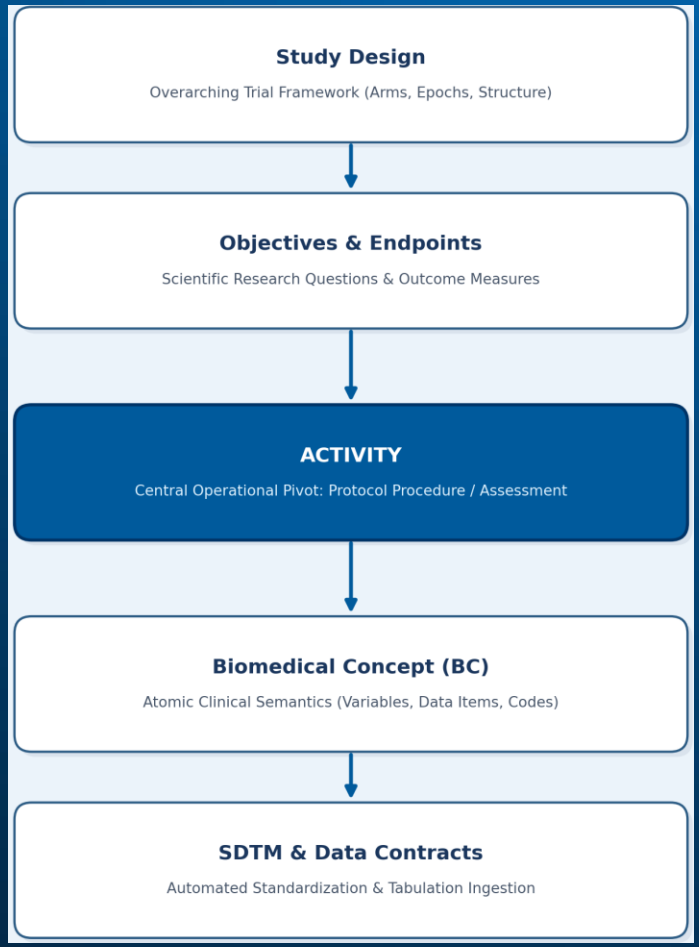
Objectives & Endpoints: The True "North Star" of Digital Protocols

Digitization often starts at the Schedule of Activities (SOA) for operational setup but Objectives & Endpoints (O&E) serve as the "North Star" guiding study logic and data collection.

Objective (The "What") : Defines primary study purpose (e.g., safety, efficacy, dose-response).

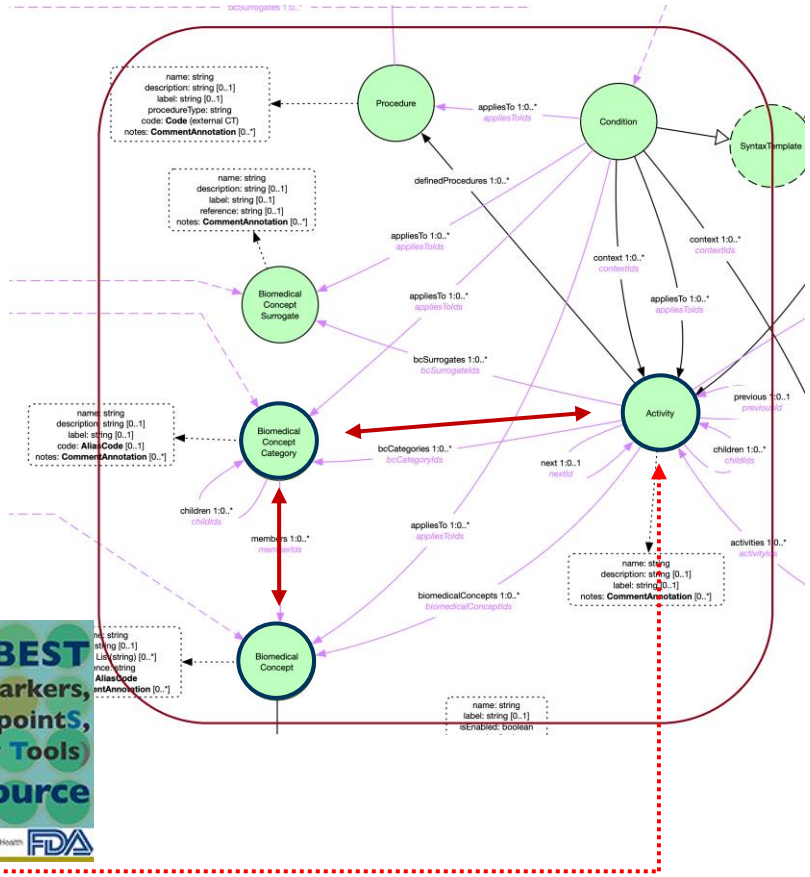
Endpoint (The "How"): Specifies measurement methods, variables, and assessment strategies (e.g., Progression-Free Survival via RECIST 1.1).

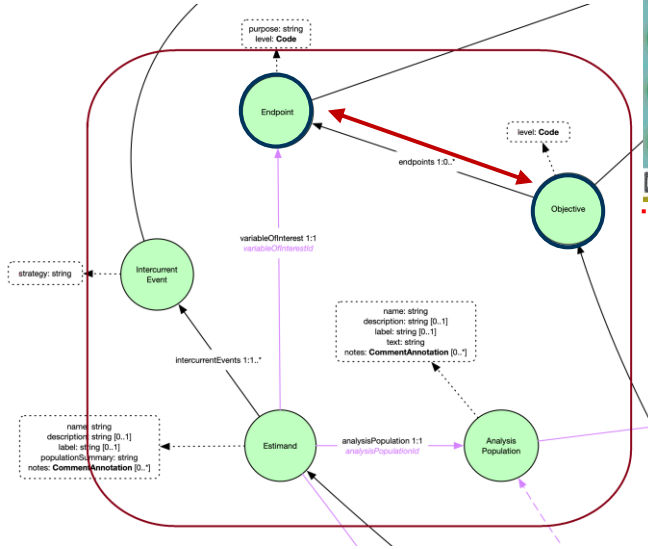
SOA Assessment (The "When & Where"): Outlines timing and location for executing procedures (e.g., blood draws, CT scans, questionnaires).

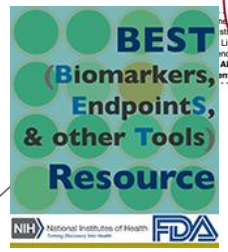


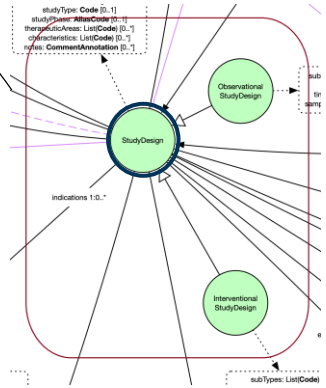
Bridging Protocol Design to Analytics

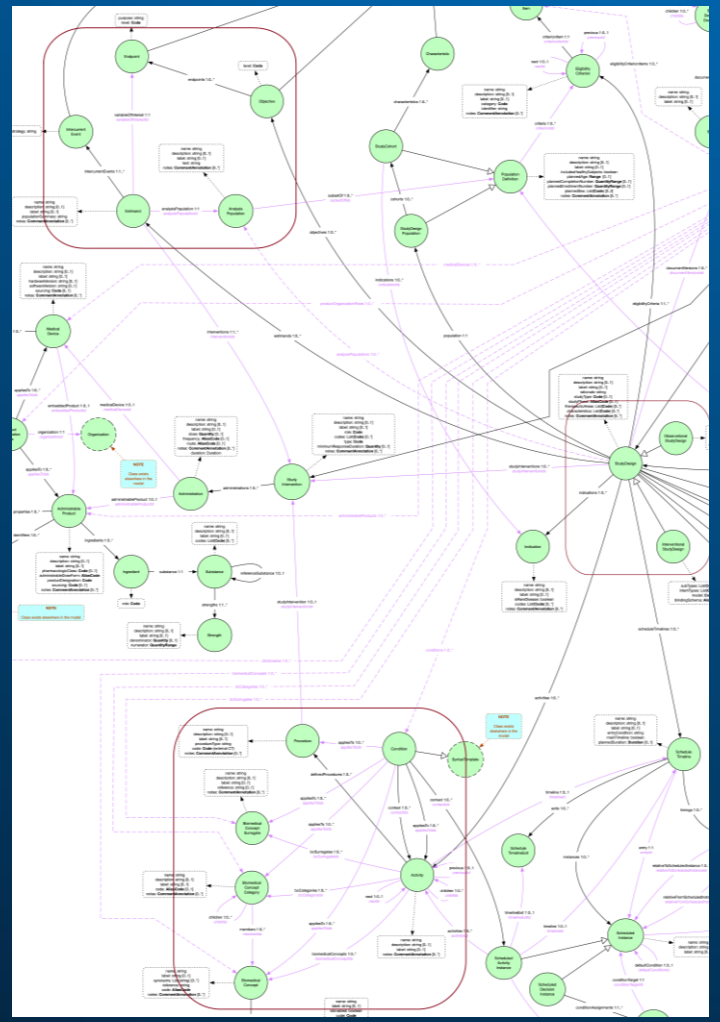
Standardized O&Es serve as the semantic bridge between protocol layout models (USDM) and downstream statistical standards (ARS/ADaM).











Endpoint Classification

Clinical Endpoints:

Directly measure how a patient **feels, functions, or survives**, evaluated using ***Clinical Outcome Assessments*** (COAs).

Surrogate Endpoints:

Indirect measures used to predict clinical benefit when direct measures are impractical or unethical (e.g., using intraocular pressure as a proxy for vision loss (glaucoma), instead of assessing the duration to go blind).

Clinical outcomes may also take many years to observe, which could delay the availability of a therapy to patients who would benefit.

Surrogate endpoints help in accelerating the access to treatment when clinical outcomes are unethical.

Clinical Outcome Assessments

Codelist Code (NCI)	Type	Definition/ Note
C95401	Patient Reported Outcome	Information collected directly from a patient that typically includes subjective feelings regarding symptoms, functions in daily life, physical, mental, emotional, spiritual, and social well-being, and satisfaction with his/her health care.
C142453	Clinician-reported Outcome	An assessment from a qualified medical professional based on examination of a subject's health status.
C142620	Observer-reported Outcome	A clinical outcome assessment that is reported by someone other than the patient or a healthcare professional.
C142632	Performance Outcome	A clinical outcome assessment in which a task is performed by a subject under the administration of a healthcare professional.
C142653	~~Proxy-reported Outcome~~	Although this type is listed in the NCI Thesaurus as a child entity under COAs, it is not listed in the NIH-FDA BEST and FDA Clinical Outcome Assessment (COA) Glossary. **The FDA does not consider a proxy-reported outcome as a valid endpoint.**

*Sponsors are strongly encouraged to participate in the **FDA COA Qualification Program**. Utilizing pre-qualified instruments drastically reduces review friction.*

Study teams should regularly consult the FDA's biannually updated list of qualified COAs to leverage previously vetted instruments or confirm acceptance criteria early in protocol design.

Surrogate Endpoints

Codelist Code (NCI)	Type	Definition
C210579	Candidate Surrogate Endpoint	Endpoints backed by a strong biological or mechanistic rationale that are actively under evaluation, but currently lack sufficient clinical data across populations to definitively prove their predictive value.
C210580	Reasonably Likely Surrogate Endpoint	Endpoints with strong mechanistic and epidemiological rationale expected to predict clinical benefit, but without full clinical confirmation across all settings.
C210581	Validated Surrogate Endpoint	Endpoints supported by extensive clinical data and clear mechanistic proof demonstrating that an effect on the surrogate reliably predicts a specific clinical outcome.

Reasonably Likely Surrogate Endpoint* often** *receives* ***Accelerated Approval*, requiring post marketing confirmatory trials to confirm true clinical benefit.*

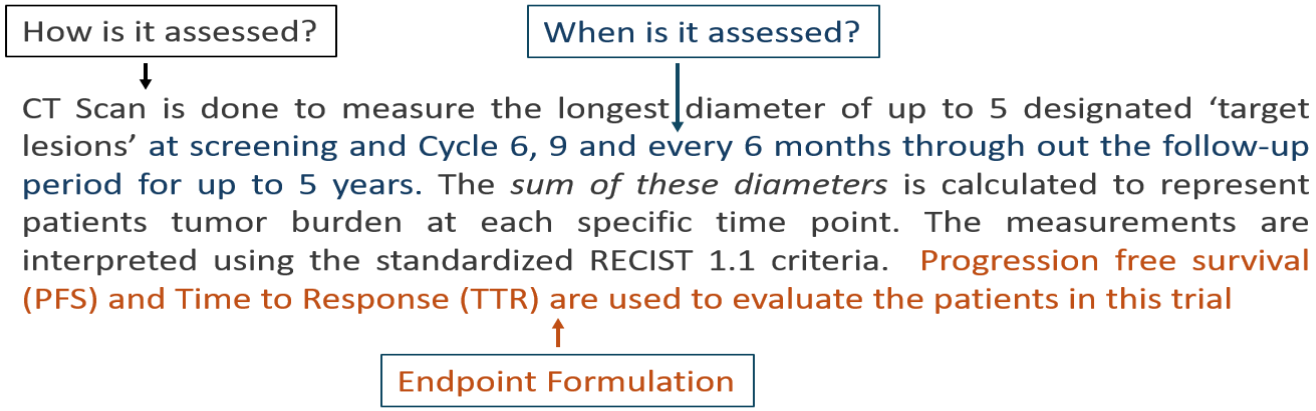
Validated Surrogate Endpoint *is supported through* ***Traditional Approval****; does not require post marketing confirmatory trials.*

Anatomy of an Endpoint

Assessment System (How): The measurement tool, target parameter, and clinical interpretation criteria used (e.g., CT scan + lesion diameter + RECIST 1.1).

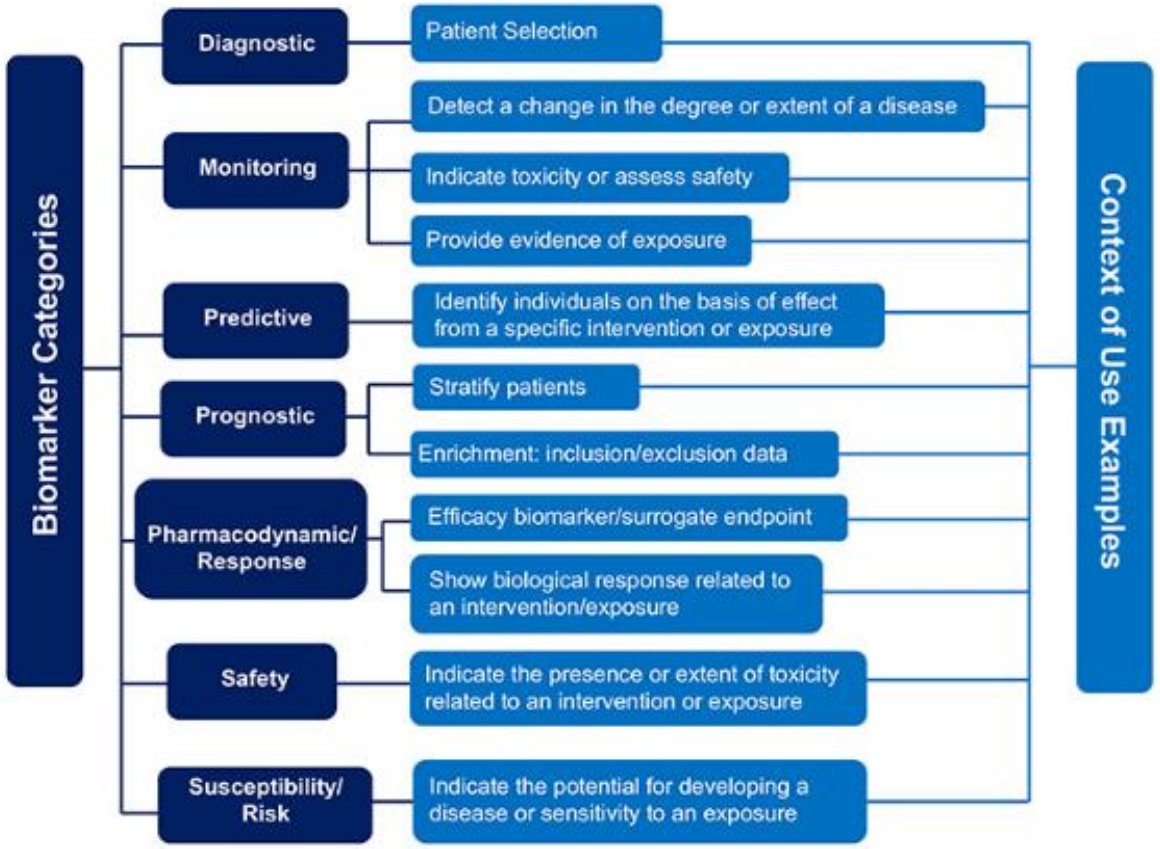
Temporal Schedule (When): The visit timepoints, baseline anchors, and total observation duration.

Endpoint Formulation (Math): The derived statistical variable analyzed (e.g., Change from Baseline, Time-to-Event, Incidence Rate).



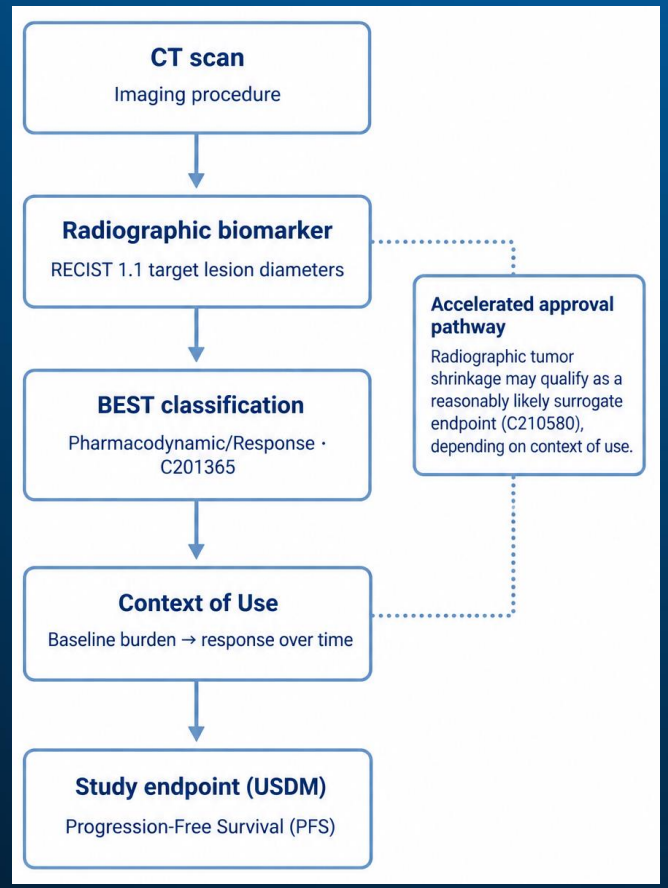
Extracted Component	USDM Entity
CT Scan	Procedure [Imaging Modality]
Target Lesions & Longest Diameters	Biomedical Concepts [Anatomical Site & Measurement Parameter]
Cycle 6, Cycle 9, q6m	Timing [Study Relative Day & Frequency]
PFS	Analysis Concept
Time to Response (TTR)	Estimand [Derivation rules-> AnalysisResultsStandard]

Biomarkers and Context of Use

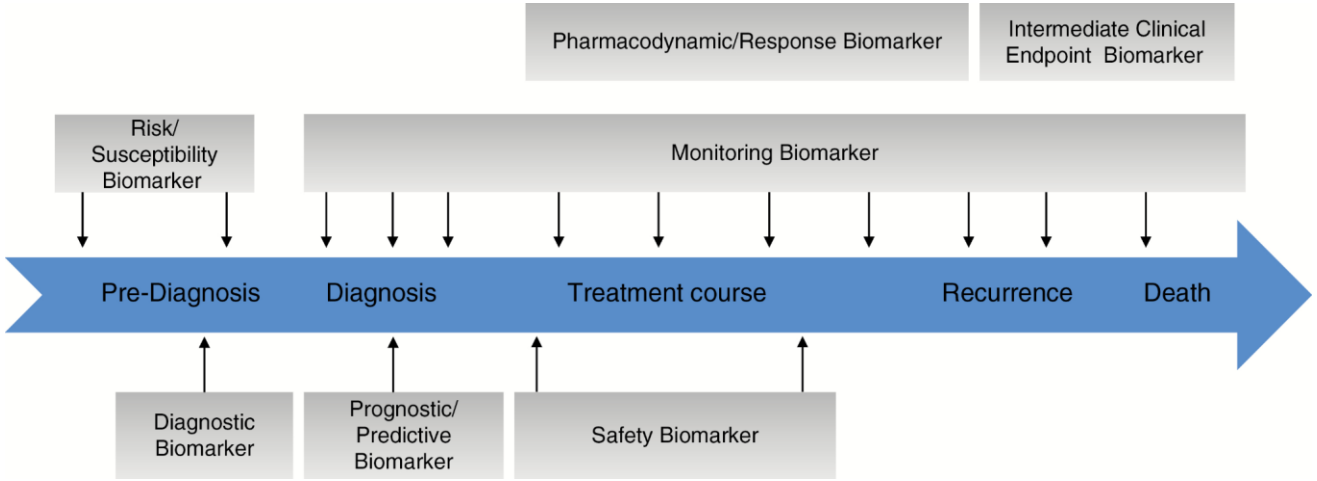


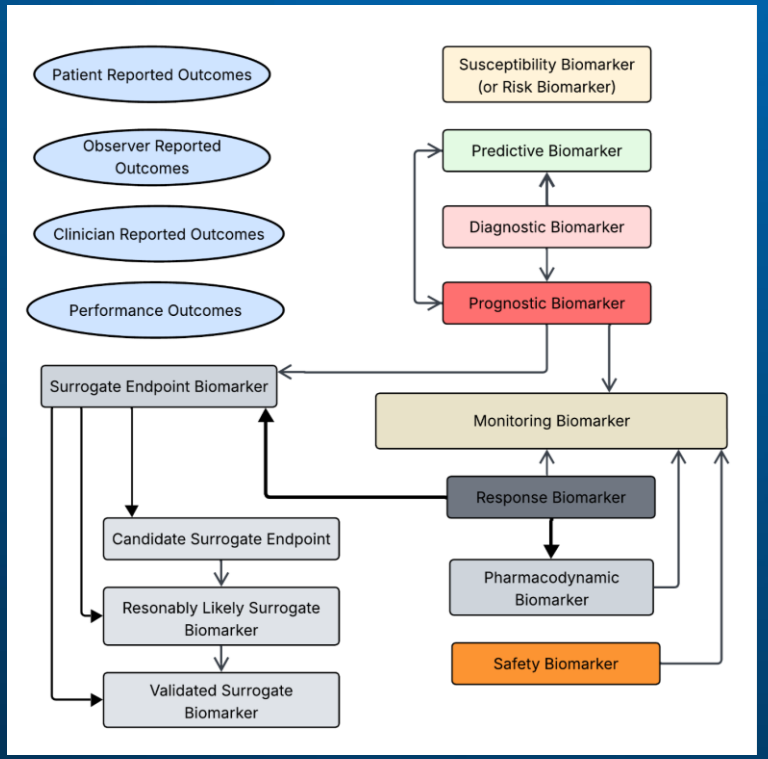
Biomarker *is defined as an objective biological characteristic (molecular, histologic, radiographic, or physiologic) measured as an indicator of biological or pathogenic processes.*

A biomarker is defined entirely by its ***Context of Use (COU)***



Biomarkers connect the dots of Patient Journey





Breast Cancer as an Example

BEST Biomarker Category	NCI Concept Code	Breast Cancer Example and Context of Use (COU)
Susceptibility / Risk Biomarker	C201360	***Germline BRCA1/2 variant:*** identify people without breast cancer who have increased inherited risk for a breast cancer prevention or screening study.
Diagnostic Biomarker	C201361	***ER, PR, and HER2 status on a diagnostic breast tumor biopsy:*** classify participants into receptor-defined breast cancer subtypes at study entry.
Prognostic Biomarker	C201363	***Oncotype DX Recurrence Score in eligible early HR-positive, HER2-negative breast cancer:*** group participants by estimated recurrence risk for a defined trial analysis.
Predictive Biomarker	C201364	**HER2-positive tumor status:** select patients with HER2-positive breast cancer for a trial of HER2-targeted treatment.
Pharmacodynamic / Response Biomarker	C201365	**Change in measured breast tumor size on MRI after neoadjuvant treatment:** assess whether the tumor is shrinking during a defined treatment study.
Monitoring Biomarker	C201362	**Repeated measurements of a known tumor on protocol-scheduled scans during treatment for metastatic breast cancer:** track disease status over time in the trial; not routine surveillance of asymptomatic survivors.
Safety Biomarker	C201367	**Left ventricular ejection fraction (LVEF) measured before and during trastuzumab treatment:** detect a decline in heart function and inform treatment safety decisions.

Evidence-Based Standardization

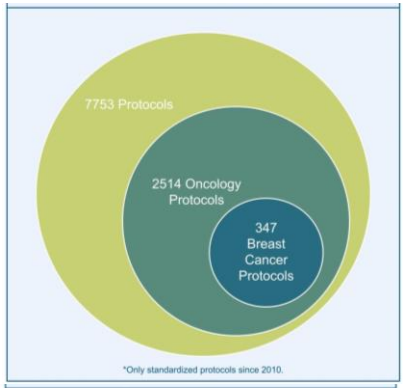
Key Quantitative Findings:

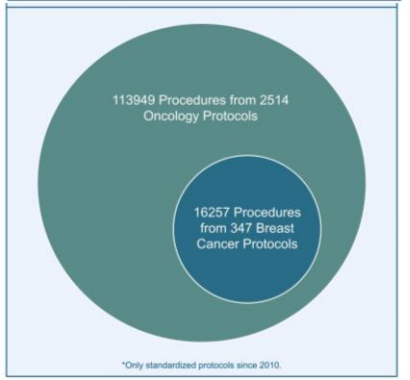
- **Repository Scope:** 7,753 total standardized protocols analyzed since 2010.
- **Oncology Core:** 2,514 oncology protocols generating 113,949 total procedure instances.
- **Breast Cancer Subset:** 347 protocols generating 16,257 total procedure instances.

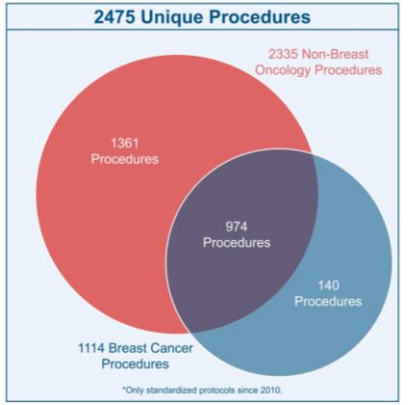
The Breast Cancer Benchmark (Venn Breakdown):

- **Shared Core Procedures** (87.4% Overlap): Breast cancer protocols draw heavily on a universal oncology baseline (e.g., Her-2/neu, CA 15-3, Diagnostic Mammography).
- **Indication-Unique Procedures** (12.6% Delta): Specialized procedures unique to breast cancer (e.g., Breast Ultrasound, Estrogen Receptor Assay).

Breast cancer serves as the initial benchmark; ongoing research will quantify pairwise overlap across additional oncology sub-indications.

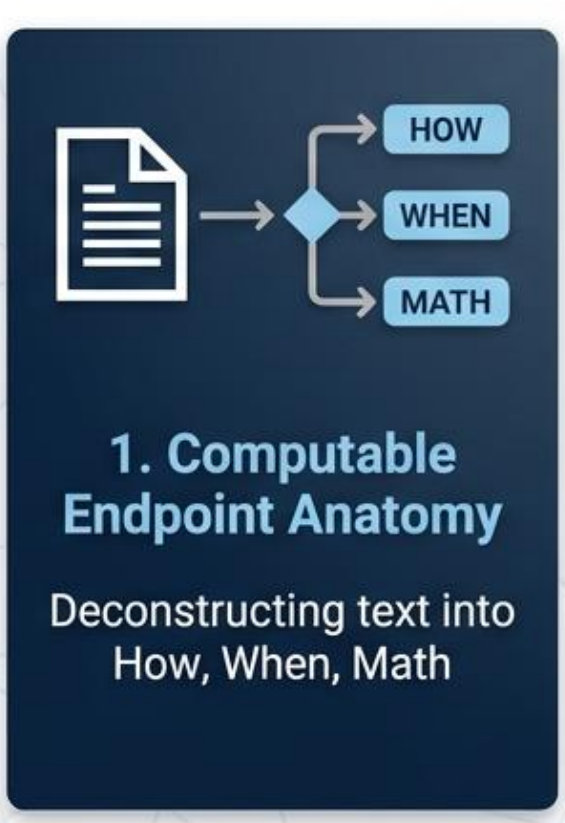






Conclusion

Four Strategic Pillars of Digital Protocol Transformation



1. Computable Endpoint Anatomy

Deconstructing text into How, When, Math



2. Context of Use Alignment

Binding BEST Taxonomy to USDM Metadata



3. Modular Procedural Reusability

87.4% Shared Oncology Core + 12.6% Specialty Extension



4. End-to-End Digital Flow

Connecting Protocol Writing to SOA Auditing and ARS Statistical Automation

Thank You!

Appreciate the contributions of Rafael Campo, Johanna Bronner, Mark Ediger and the entire Data Curation Team and other cross-functional teams.

