SCDM Live

Sept. 29 – Oct. 2, 2024 | Boston, MA

SCDM 2024 Annual Conference

The Festival Of Opportunity

Baltimore, MD - February 7, 1904

Using Standards to Make Clinical Research More Effective

Chaired by

Phil Kirsch

Director of Quality

DF/Net Research, Inc.

Joseph Lengfellner

Senior Director, Clinical Research Informatics

Memorial Sloan Kettering Cancer Center

Drew Garty

Chief Technology Officer, Clinical Data

Veeva Systems

Standards changed the world in the 20th Century

- 150 years ago, everything was made by hand.

- Then, Henry Ford pioneered the assembly line.

- And the US Department of Commerce began creating standards.

- In 1900, the average US home had
  - Icebox
  - Sewing Machine
  - Washing Machine
  - Stove
  - Clock

- Today, the number is 25-30

How does standardization make a difference?

- Early cars cost between $1500 and $3000 – two or three times more than a typical house. The first Model T sold in 1908 for $850; by 1925, it was down to $300.

- The British had been using 55° screw threads since 1840, but during WW II, they switched to the US Standard (60°) to use American War Planes.

- Early computers were created for unique purposes. Today, we use Word Processors, Spreadsheets, etc., but it would be possible without operating systems and programming languages.

But Clinical Trials remain Unique, Hand-Crafted Items

- Most trials contain more unique CRFs than standard forms.

- Many clinical databases are built one field or module at a time.

- Every company has its own format for study protocols.

- CDISC standards still allow a significant amount of customization.

- Electronic Health Records aren't readily usable for clinical research.

All Clinical Protocols include …

- Measurable Endpoints
- Schedule
  - Visit
  - Treatment
  - Follow-Up
- Inclusion and Exclusion Criteria

Many include

- Toxicity Tables
- Statistical Analysis Plans

So, an intelligent system should be able to

- Design a database capable of capturing study data.

- Identify expected consistency checks.

- Build a sample database based on each element's normal distribution.

- Include unexpected values to test legal ranges and verify that the system flags missing or conflicting values.

- Actively compare actual data to those expectations during a trial and flag significant deviations for human review.

A simple example

Systolic Blood Pressure:
•**Mean**: 119.86 mm Hg
•**Median**: 120.10 mm Hg
•**Mode**: 61.95 mm Hg
•**Standard Deviation**: 15.06 mm Hg

Diastolic Blood Pressure:
•**Mean**: 80.05 mm Hg
•**Median**: 80.02 mm Hg
•**Mode**: 45.91 mm Hg
•**Standard Deviation**: 9.97 mm Hg

Distribution of Systolic Blood Pressure

Distribution of Diastolic Blood Pressure
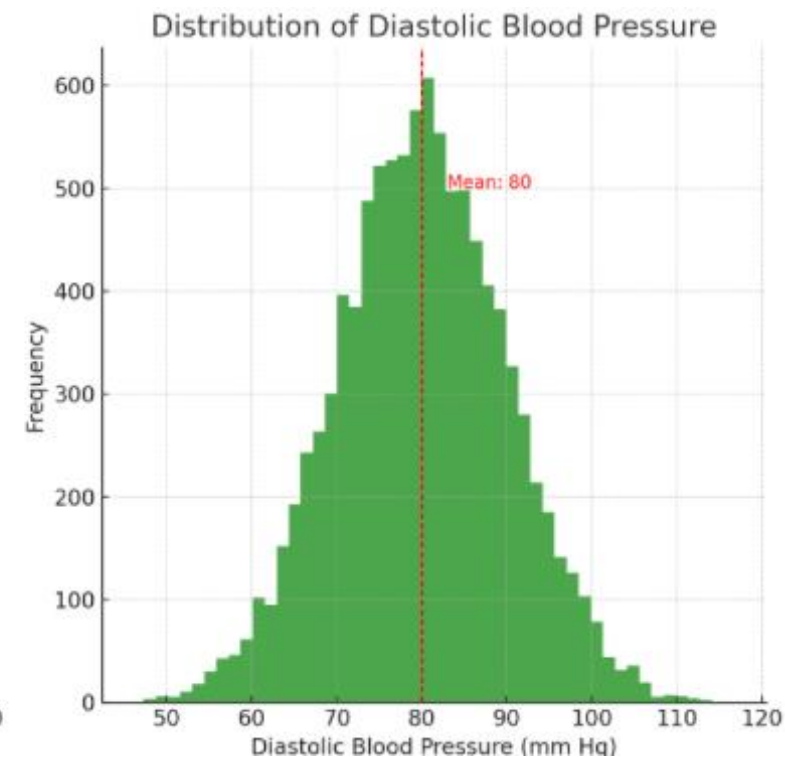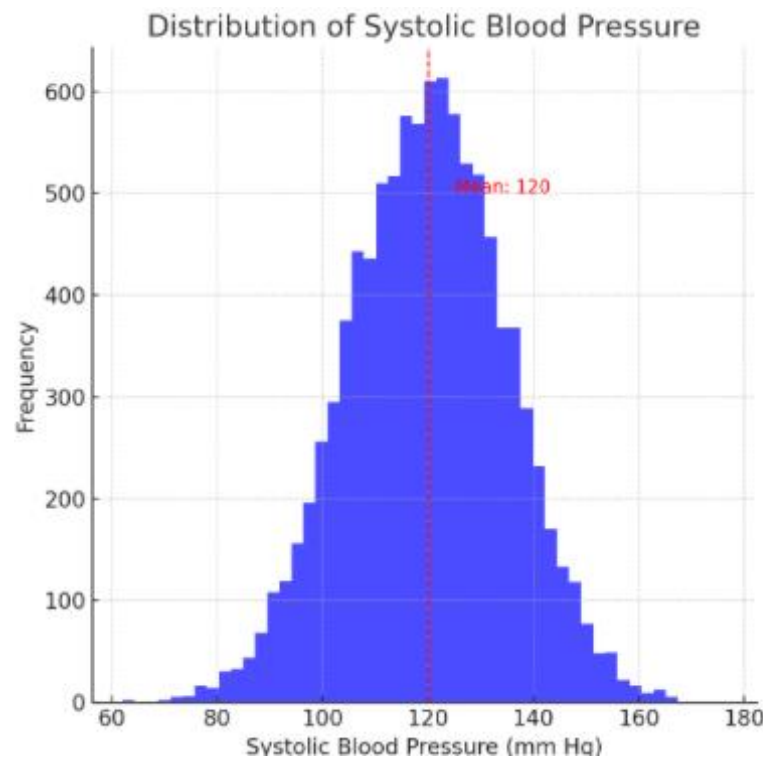
Objections

- History
  - We've never done it that way before
  - It won't meet our unique needs.

- Cost

- Validation

Responses

Neither had Henry Ford.

One black model of one product resulted in the variety of manufactured products available today.

Initial Investment will pay off quickly.

Individual databases could still be validated and approved.

New standards could streamline validation.

The whole point of clinical research is to compare treatment results across a representative sample of the target population.

New standards could speed database development

- Clinical Protocols could be required to include certain XML or JSON tags …
  - Title
  - Primary Endpoint
  - Secondary Endpoint
  - Inclusion Criteria (#)
  - Exclusion Criteria (#)

- Clinical Databases could then be programmed to automatically configure databases to capture the elements needed.

Standardized use of existing data can increase quality and lower the cost of review.

- An international catalog of human data ranges could be a source for clinical scientists worldwide.

- Computerized systems can flag significant deviations for further human review.

Standards can reduce costs
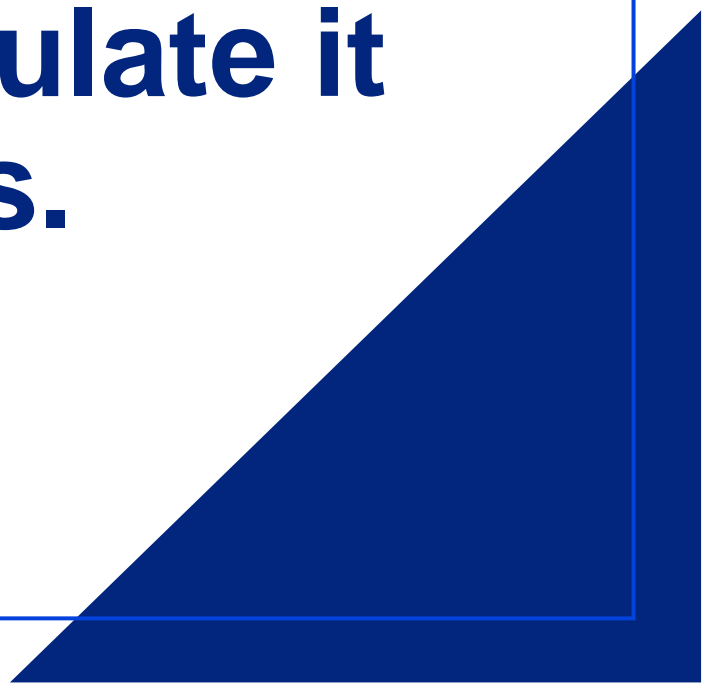
- ## Token Cost

  An LLM requires more "tokens" and time to read a study database and select the elements needed to answer a question than if a human directs it to key elements.

- ## Hallucination Risk

  By focusing on critical elements, we can reduce the likelihood of the computer including irrelevant considerations.

If computerized systems can pass the MCAT, GRE, Bar, etc., they should be able to build a clinical database and populate it with typical sample values.

Expanding Standardization and Making it Effective

- Joe Lengfellner – Sr. Director, Clinical Research Informatics

Memorial Sloan Kettering
Cancer Center

The Mounting Challenges Faced By Provider Sites Today

"What we're seeing is the consequence of biopharmaceutical companies engaging in more ambitious and customized drug development activity that targets a growing number of rare diseases, stratifies participant subgroups using biomarker and genetic data, and relies on more structured and unstructured patient data from a larger number of sources,"

Ken Getz, Professor and Director of Tufts CSDD

MSK's Experience Mirrors Larger Trends and Challenges:

- Increased number of smaller, targeted trials

- Trial Complexity is Increasing
    - Increased data requirements from sponsors
    - Increased queries from sponsors/CROs
    - More patient management activities

Active Treatment Patients on Study

Year	Active Treatment Patients
2019	12648
2020	12762
2021	14637
2022	16184
June 2023	16683
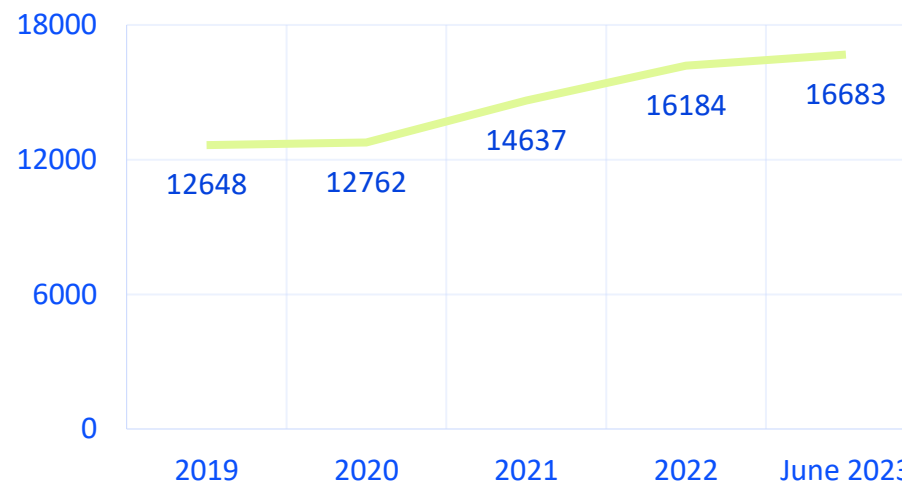
Persona

Claire,
The CRC

"Data entry often feels like a clinical decision."

Experience level
1.5 years

Key Characteristics

- **In charge of abstracting patient data from MSK tools to input into several sponsor EDCs**
- Works on **4 - 8 protocols simultaneously**
- **Interacts with CRAs/RPAs, CTNs, PIs,** Sponsor Monitors and Auditors
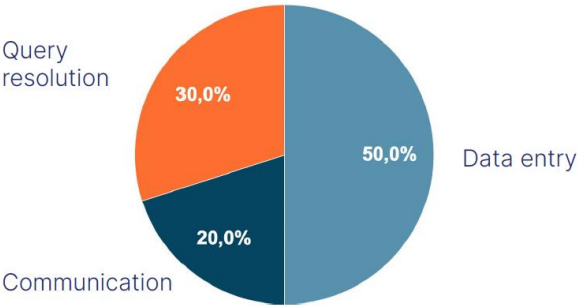- Usually works with double screen and self-organizes her workflow and follow up the data-entry process.

Goals

- Enter patient data in an **accurate and timely manner**
- **Solve queries** as quick as possible
- **Effectively communicate** with other professionals to gather the information that is missing

Needs

- **Normalized patient information** (right after visit)
- **100% populated fields** in both HIS and EMR
- Clear query instructions
- Time-saving hacks or shortcuts for recurrent data entry fields
- Quick, effective communication channels

Time Allocation (Weekly)

Query resolution 30,0%

Data entry 50,0%

Communication 20,0%

Tools

Memorial Sloan Kettering Cancer Center

- **EMR** / HIS
- **CIS**
- CARS
- Teams:
  - Shared Drive
  - Email
  - Calendar

Google

Veeva

Sponsor EDC&CTMS

medidata

Datalabs EDC

REDCap

viedoc

ORACLE Health Sciences

Pfizer

ICON   iSite

38

Mats Sundgren, PhD
Industry Science Director, i-HD

Gynet Santiago
Clinical Research Coordinator, Memorial Sloan Kettering Cancer Center (MSK)

Joseph Lengfellner
Sr. Director

Streamlining Clinical Trials with eSource: Insights from MSK

Memorial Sloan Kettering use case explores the potential of EHR-to-EDC.

Today, clinical trial data collection faces increased complexity due to data duplication between research systems and electronic health records (EHRs), particularly in oncology studies. Manual processes, such as data entry, consume significant time and resources. Over 50% of clinical trial data is duplicated between research systems and hospital EHRs, with around 20% of total study costs typically allocated to data duplication and verification.[1]

The number of data points collected in oncology trials has surged dramatically over the past decade, driven by advances in genomic technologies, digital health tools, and personalized medicine. While this increase in data volume enhances our ability to understand and treat cancer more effectively, it also presents significant challenges in data management and analysis. In typical Phase III oncology trials, approximately 10,000 data points are collected per patient. However, in trials incorporating genomic and digital health data, this number can multiply ten to a hundredfold.[2,3]

As a result, study sites, particularly clinical research coordinators (CRCs), face increased workloads, complex data management tasks, and the need for specialized training, leading to higher personnel costs. Sponsor companies incur additional expenses due to investments in advanced data management tools, extended trial timelines, and source data verification (SDV). To mitigate these challenges, sites and sponsors must explore innovative solutions, collaborate with stakeholders and vendors, and ensure regulatory compliance to improve efficiency, reduce costs, and enhance cancer treatment outcomes.

some nature of current double-entry processes and supports the three-minute-per-data-point estimate. Additionally, it revisits the value generated for investigational sites and sponsors by introducing this technology, exploring the transformative potential of eSource for future clinical trials.

Methodology

The study employs a qualitative structured interview guide to understand the data entry process during clinical trials. An in-depth, semi-structured interview was conducted with an experienced CRC to gather insights into workflows, challenges, and perceptions regarding data entry tasks. Open-ended questions allowed for rich, nuanced responses, capturing the complexity and variability of the process. The validity and generalizability of the interview outcomes were assessed and found to correspond well with CRC practices at MSK.

CRCs: Key players in trial success

Gynet Santiago is a seasoned CRC at MSK, with 16 years of experience and involvement in over 50 clinical studies. She summarizes her role as, "I am responsible for collecting, extracting, and entering data for research projects, databases, and clinical trial protocols. This includes reviewing patient charts, existing databases, and other sources within a specific timeframe."

The CRC is pivotal in executing clinical trials, with responsibilities spanning various trial phases to ensure compliance with regulatory requirements and ethical standards, while prioritizing patient safety

Standardization at Sites - Why is it important?

- **Diverse portfolio**
  - At MSK, 2000 research studies with 400+ sponsors, 400 research coordinators

- **Turnover in research coordinator roles**
  - Nearly 30% turnover in clinical research professionals[1]

- **Tool Creation Significant portion of startup timeline**
  - Average startup time at AMC – 8.12 months[2]
  - Consent form, treatment order sets, CTMS calendar setups

- **Data Management Costs**
  - Data management expenses can account for 25% of overall clinical trial budget[3]

1) https://www.appliedclinicaltrialsonline.com/view/salary-survey-healthy-trajectory-clinical-research-professionals
2) https://www.wcgclinical.com/wp-content/uploads/2024/01/wcg-trends-insights-2024.pdf
3) https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4386950/

Examples of Non-Standardization

- **Multiple technology solutions**
  - Software
  - Hardware (i.e. EKG machines)

- **Inconsistent login credentials**
  - 8-12 systems/study

- **Differences in data requirements from sponsors**
  - Different CRF design
  - Inconsistent data management plans

- **Variable study eligibility criteria**
  - Subtle differences in criteria
  - No consistency makes automated matching harder

Thoughts on a Better Future

- **Technology solutions that fit into existing site workflows**

- **Improved industry/site collaboration**
  - Standard protocol design
  - Enabling technology such as EHR-to-EDC

- **Enable sharing of tools between sites**
  - Treatment order set builds
  - CTMS builds

- **Continue advocation for standardizing data interoperability standards (i.e. FHIR)**

- **Look for opportunities to bring sites into industry discussions**

Memorial Sloan Kettering
Cancer Center

Simplifying Standards to Streamline Studies

Drew Garty, CTO, Clinical Data

The Data Source Tipping Point
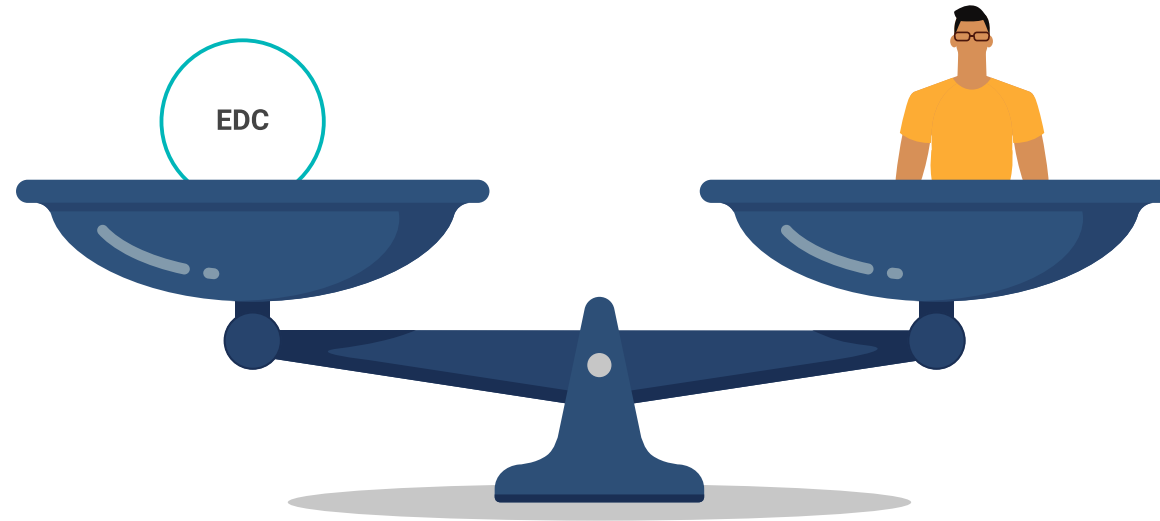
The Data Source Tipping Point

The Data Source Tipping Point

SCDM Live

EDC

Imaging Results

Labs

ePRO

RTSM
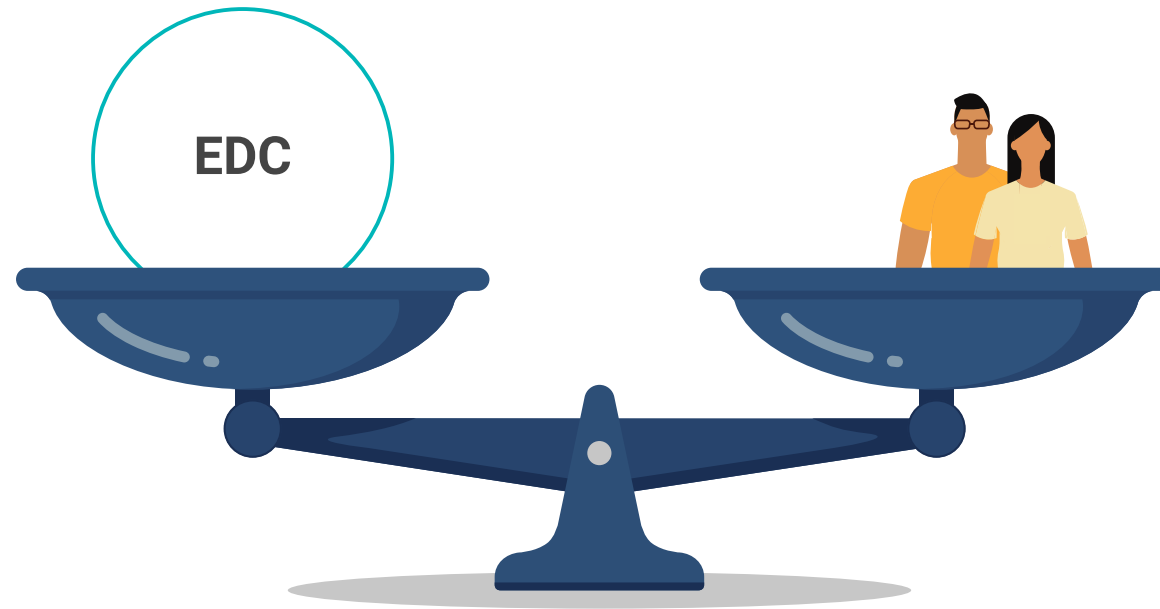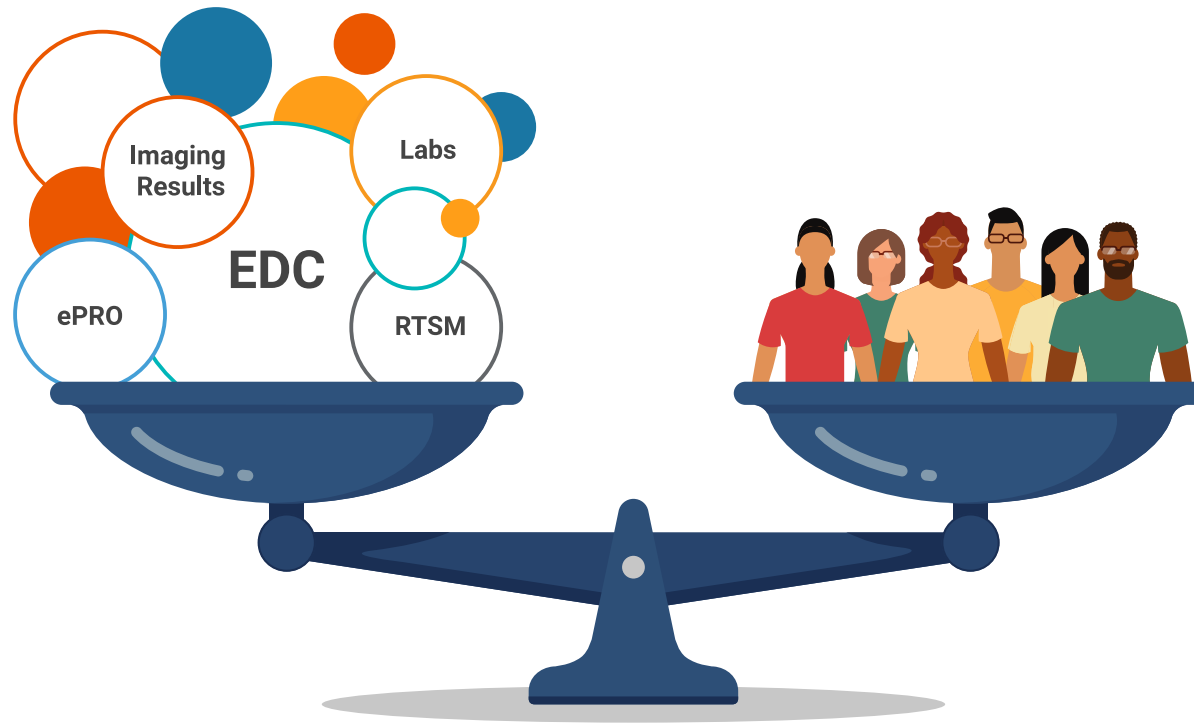
The Data Source Tipping Point
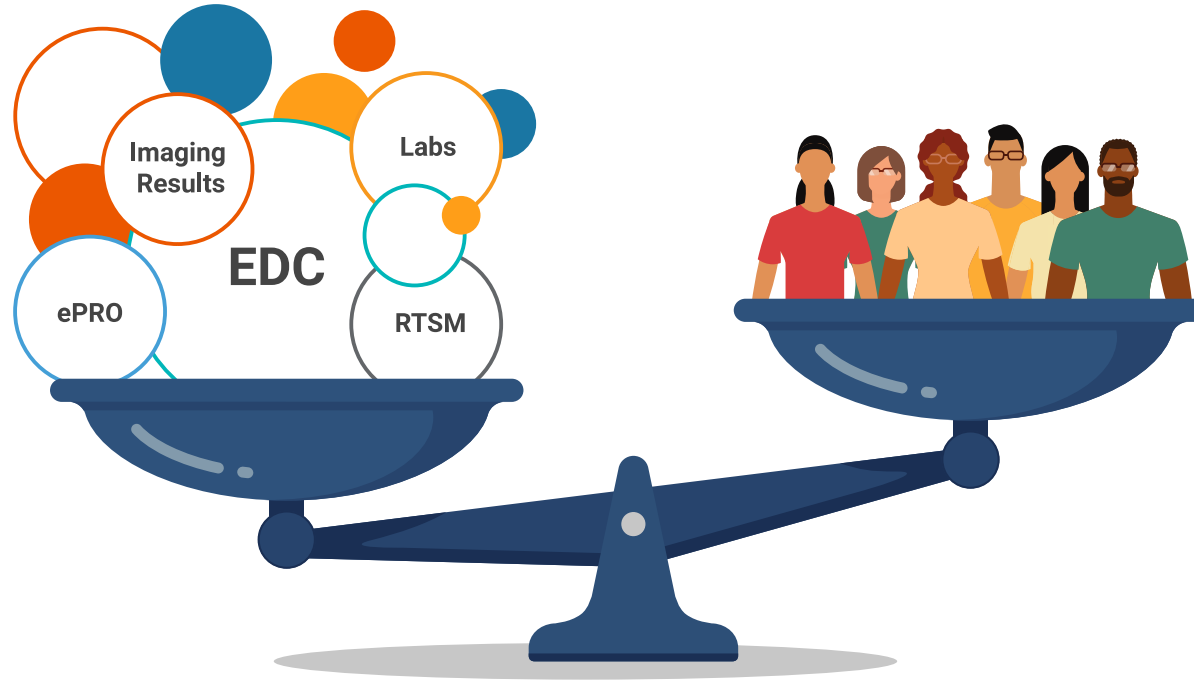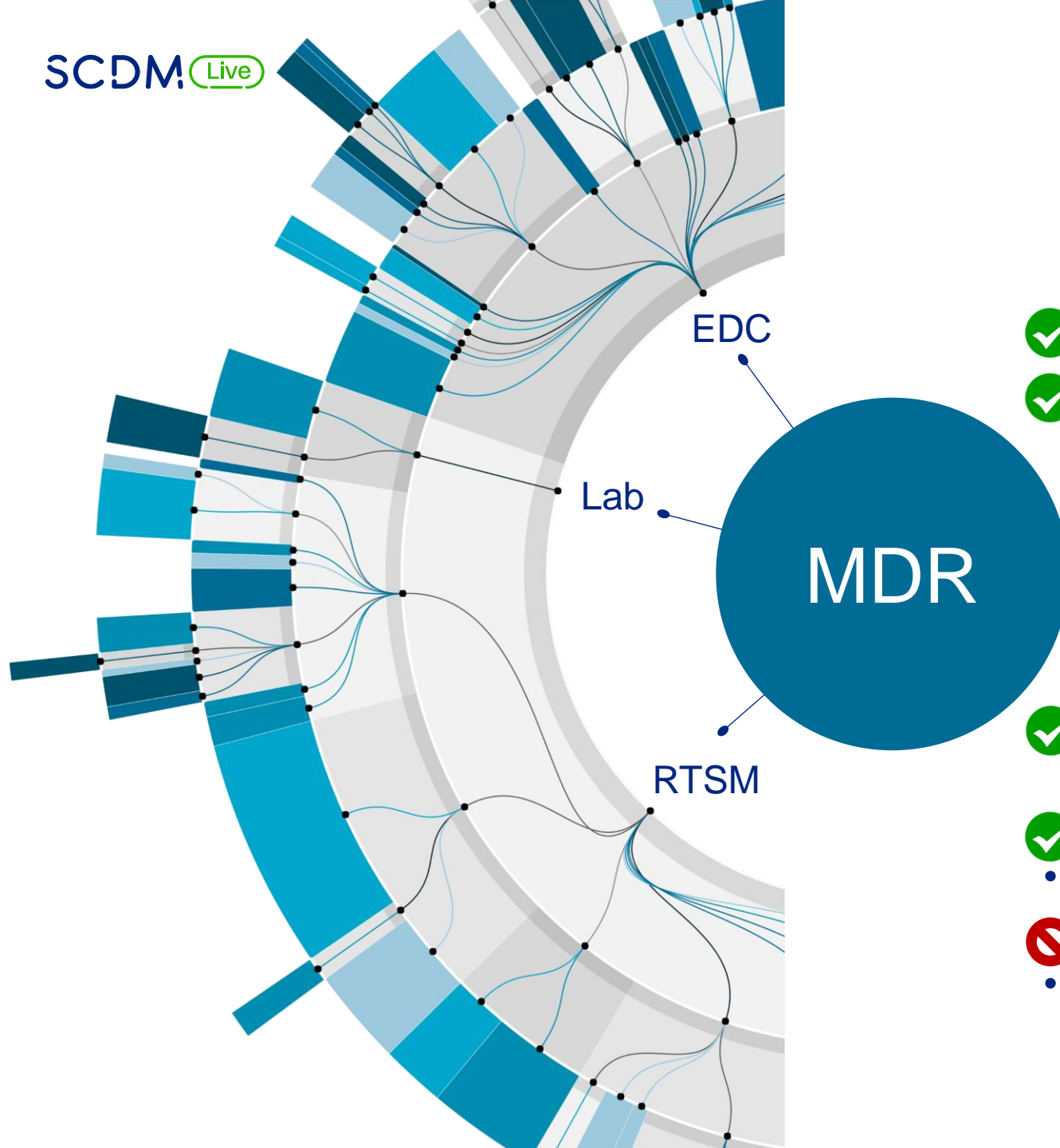
Today's Leading MDR Strategy

✅ Design studies with the end in mind (analysis)

✅ MDR is an important solution to tie together:

- study design
- data collection
- analysis
- submission

✅ Study design should start with MDR and through relationships define data collection and analysis

✅ Data Collection Design should be as automated as possible

🚫 MDR should be a repository of all (or as much) data collection metadata to automate collection build

EDC

Lab

MDR

RTSM

SCDM Live

Scalability of Centralized Metadata Management

- How many unique sources of data today and in the future?

- How many active versions of those systems?

- How many unique properties/attributes per source system?

- How many property settings per attribute per study?

- What efficiency do you lose in study build by working outside of the system (display logic/rules/etc)?

- What fragility is created by working outside the data collection system off the shelf design environment?

- What is the cycle time that will be necessary to add or more importantly modify external metadata?

Common Metadata Between Data Collection and Data Analysis

Data Collection	Data Analysis

Terminology For This Discussion

Data Collection | Data Analysis

Data Definition

Specific study design metadata that is shared, common, and critical to both DM and Stats

Terminology For This Discussion

Data Definition	Data Standard	Data Collection Standard
Specific study design metadata that is shared, common, and critical to both DM and Stats	Agreed upon standardization of data definitions	additional metadata built upon the Data Standard to maximize value of reuse of EDC or other data collection tool

Goal: Enabling Stats and DM Parallel Flow

Stats

Annotations

SDTM Programming

Data Def.

Go Live

EDC CRF Design

EDC Rules and Dynamics

DM

Relative Scale of Metadata
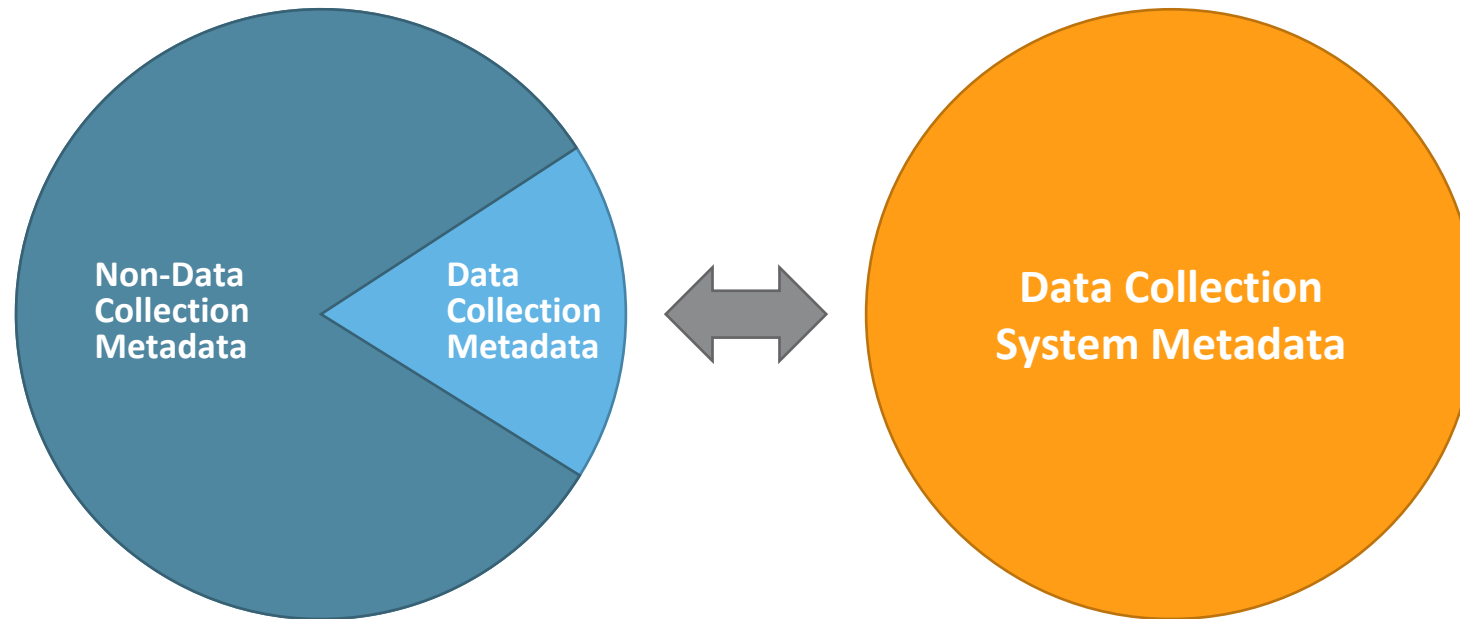
How to Automate Governance and Compliance to Data Standard and Definition?
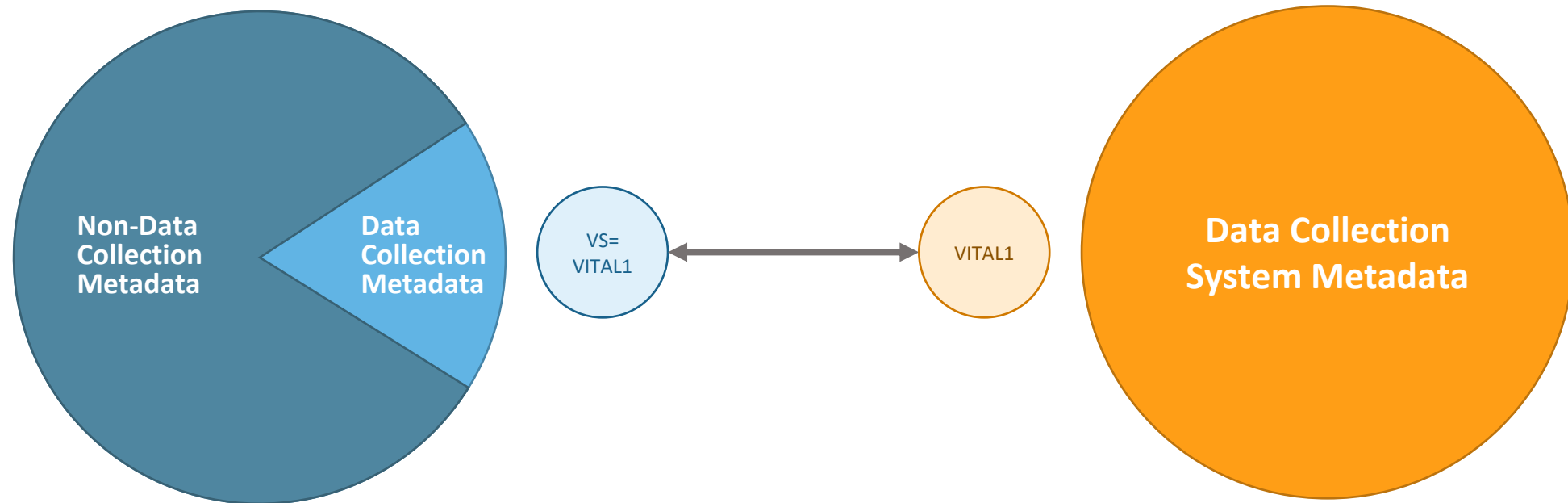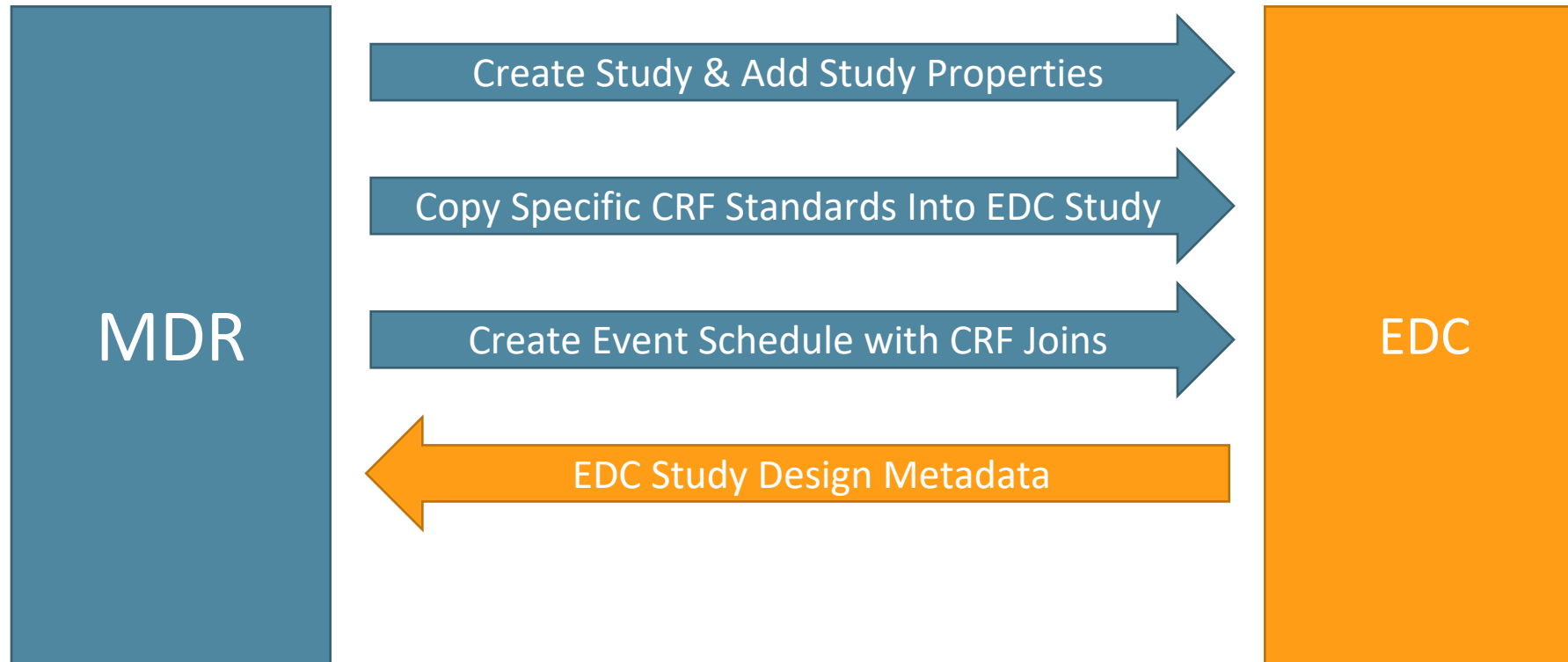
Non-Data Collection Metadata

Data Collection Metadata

VS= VITAL1

VITAL1

Data Collection System Metadata
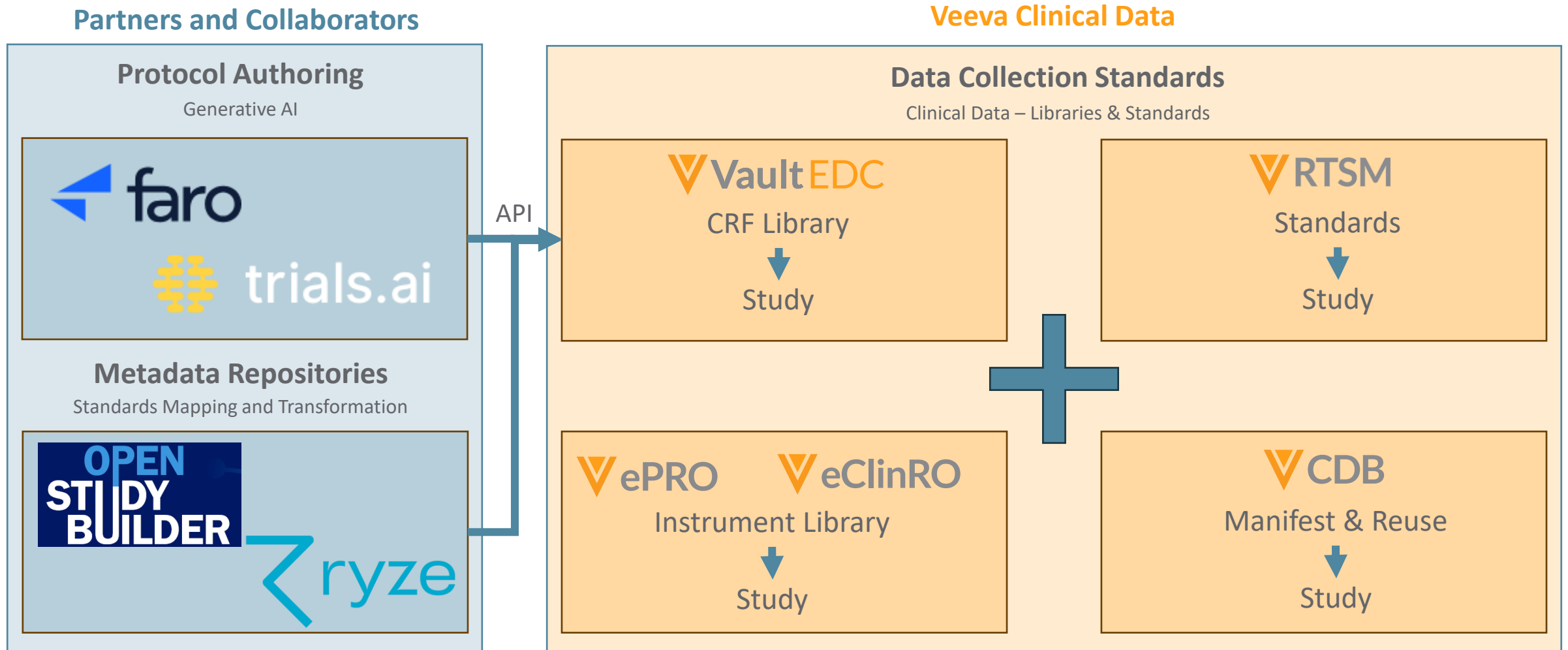
Where to go from here?
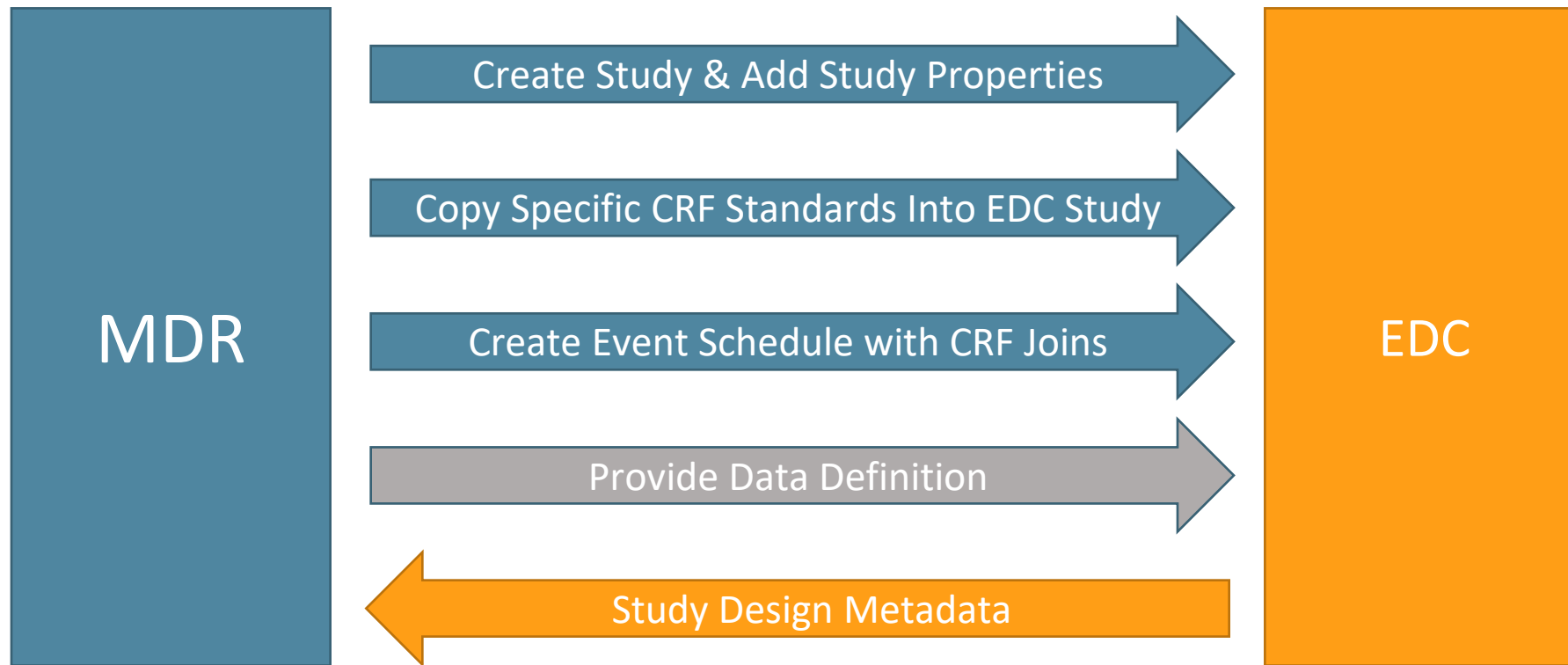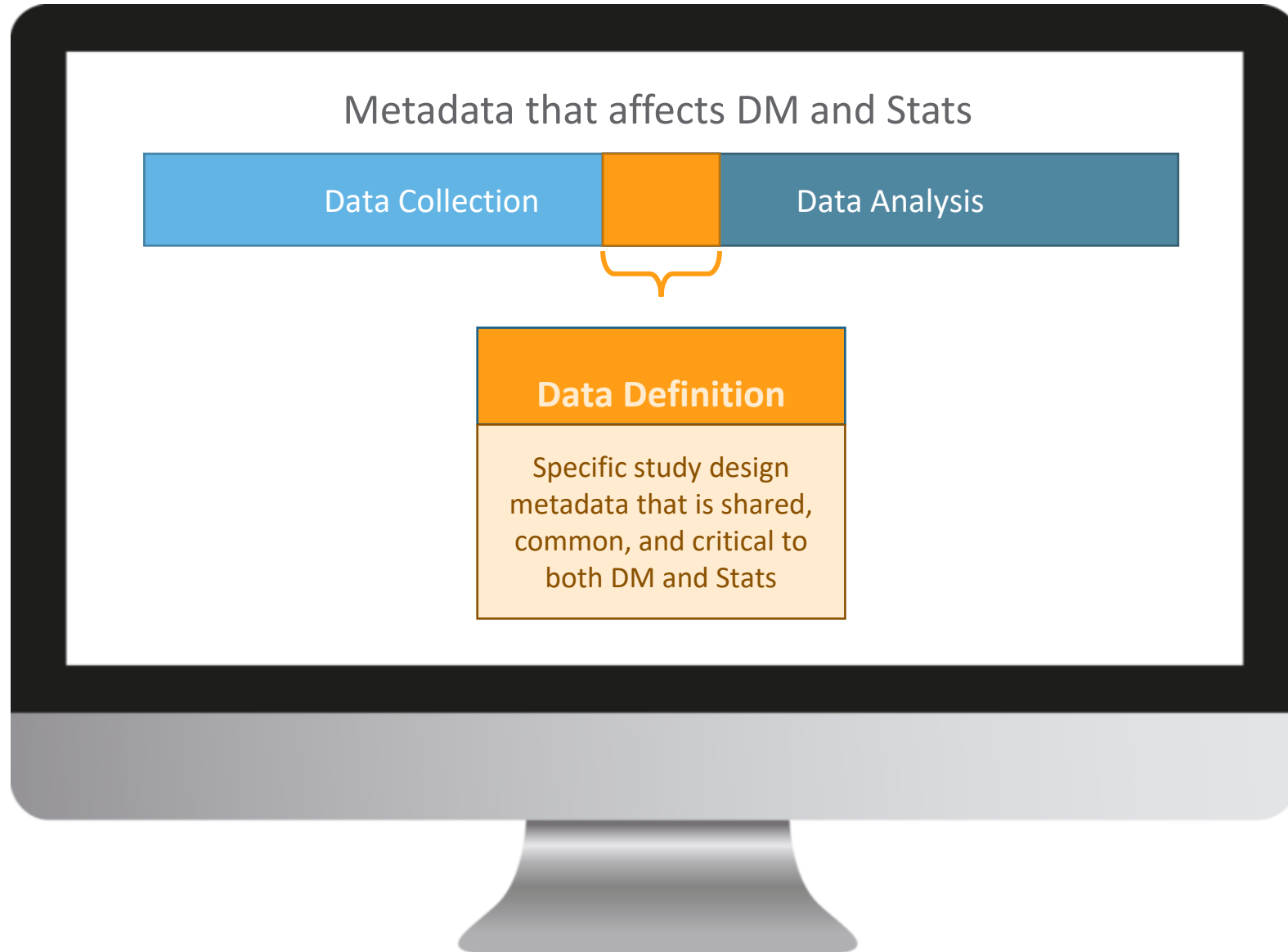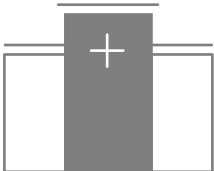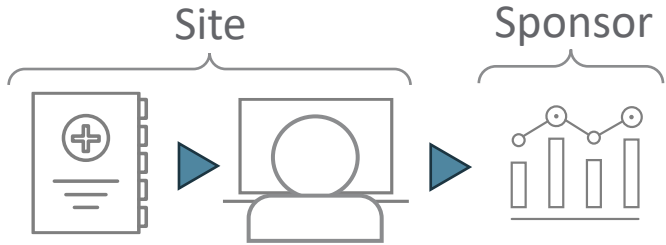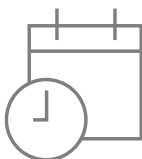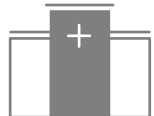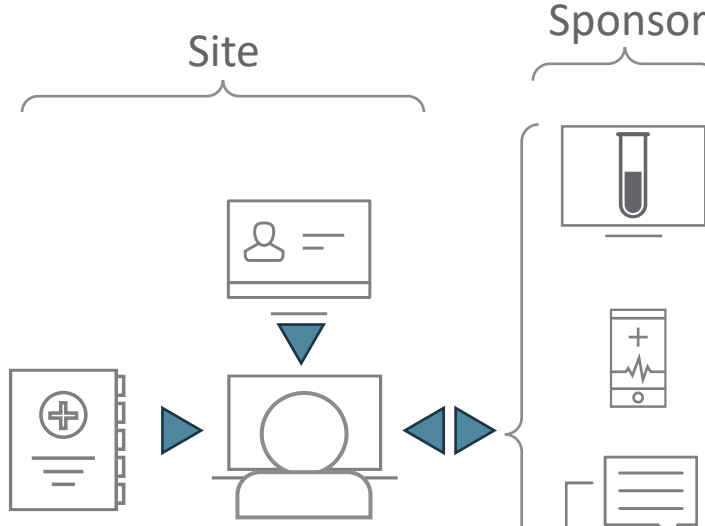
MDR

Create Study & Add Study Properties →

Copy Specific CRF Standards Into EDC Study →

Create Event Schedule with CRF Joins →

Provide Data Definition →

← Study Design Metadata

EDC

SCDM. Live

Simplifying the Site Data Experience

How we can help

A lot has changed in 30+ years: The New Site Experience

SCDM Live

	WHO?	WHERE?	WHAT?	HOW?	WHEN?
1990s				Site → Sponsor	
TODAY		Visit Method	RTSM & ePRO	Site Sponsor — Complete and Concurrent Participant Records — CDB	Automation