National Children's Study (NCS)

MDR: A Metadata Repository for Longitudinal Studies

Bethesda, MD

January 23, 2012

This document is confidential and is intended solely for the use and information of the client to whom it is addressed.

THE NATIONAL CHILDREN'S STUDY
HEALTH GROWTH ENVIRONMENT

Booz | Allen | Hamilton

Today's Agenda

- What is a longitudinal MDR?
  - A longitudinal MDR documents data collection over time
  - A longitudinal MDR documents and facilitates data management over time
- The MDR Website
- Challenges Ahead

What is a longitudinal MDR?

- A longitudinal MDR is a **database** and **website** that chronicles the conduct of longitudinal studies
  - The basic unit of a longitudinal MDR is the **data element**
  - MDR is capable of following data elements
    - from **conception**
    - through their use in **data collection**
    - through their participation in **data processing**
    - to **data dissemination**
- The MDR uses study objects and a model that conforms to international standards
  - One standard is **the Data Documentation Initiative** (DDI)
    - DDI is used both in the United States and worldwide for chronicling observational research especially with questionnaires
  - The MDR also supports the **ISO 11179** standard
    - It includes a view of data elements not unlike the view provided by the Cancer Data Standards Registry and Repository (caDSR) Common Data Elements (CDE) Browser

From the National Cancer Institute Common Data Elements (CDE) Browser

What is a longitudinal MDR?

CDE Browser

CDE Cart | Home | FormBuilder | Help

Data Element Search

Search for Data Elements

1 Matches | Search preferences | Advanced search

caDSR Contexts

- Exact phrase
- All of the words
- At least one of the words

newborn current weight

Name

Tip: This is an exact match search. To search for partial words or phrases use the * as a wildcard.

Note: Default settings exclude Test and Training Context views from the tree and certain 'non-released' Workflow and Registration statuses. Click the 'Search Preferences' link above to view or change the exclusion criteria. Search Preferences' will be reset to default settings when the 'New Search' button is clicked on the search results page or 'caDSR Context' in the Tree.

Search | Clear | New Search

Search Results

Search within results

Results fewer than expected? Check Search Preferences

[Download Data Elements to Prior Excel] [Download Data Elements to Excel] [Download Data Elements as XML] [Download CDE Browser DTDs]

Sort order : (Default) Registration Status>>Workflow Status>>Long Name [Ascending]

Add to CDE Cart | Add to CDE compare list | Compare CDEs

1 - 1 of 1

Long Name	Preferred Question Text	Owned By	Used By Context	Registration Status	Workflow Status	Public ID	Version
Assessment Newborn Current Weight java.lang.Integer		NICHD			RELEASED	2958541	1.0

From the National Centers for Biomedical Ontology (NCBO) BioPortal

What is a longitudinal MDR?

Pediatric Terminology

Terms

Jump To:

- NICHD_Pediatric_Terminology
  - Anatomic_Structure_or_System
    - Body_Part
    - Embryonic_Structure_or_System
  - Device
  - Disease_or_Disorder
  - Finding
  - Healthcare_Activity
  - NICHD_Subset
  - Organism
  - Organismal_Process
  - Person
    - Attending_Health_Care_Provider
    - Child
      - Infant
      - **Newborn**
        - Full_Term_Neonate
        - Preterm_Newborn_Infant
        - Sick_Newborn
      - Toddler
  - Property_or_Attribute
  - Substance

Details | Visualization | Notes (0) | Term Mappings (0) | Ter

Preferred Name	Newborn
ID	Newborn
Full Id	http://www.owl-ontologies.com
Definition	An infant during the first month
Nci	C16731
Nci Pt	Newborn
Sub Class Of	Child

What is a longitudinal MDR?

NCS METADATA REPOSITORY

THE NATIONAL CHILDREN'S STUDY

Logged in as Art Vandelay | Skip Nav | Help | Logoff

Home | About MDR | Data Dictionary | Change Logs | Data Element Search | Browse Study | Resource Center

Browse Study

Operational Data Elements | Instrument Data Element

Recruitment | Pre-Conception | Pregnancy | Birth | First Two years | Years 2-21

Pre-Conception is a series of collection events. During this period study participants who were not already pregnant at the time of enrollment are followed to determine if they have become pregnant. Initially the participant's Pregnancy Probability Group (PPG) is calculated. Follow-ups are scheduled based on the participant's PPG status.

Each phase of the study has one or several collection events at which data is collected. A collection event could be an interview, phone call, or a visit in which an instrument(s) is administered. For each collection event, there is a collection of instruments that are either administered at the same time or are more than one version of an instrument. Collection events and collections of instruments that are administered at each collection event are listed below. To view a list of tables within a collection of instruments, click on the link for the desired collection of instruments below.

Collection Events

Pre-Conception Visit Event Instruments

- Pre-Pregnancy Visit

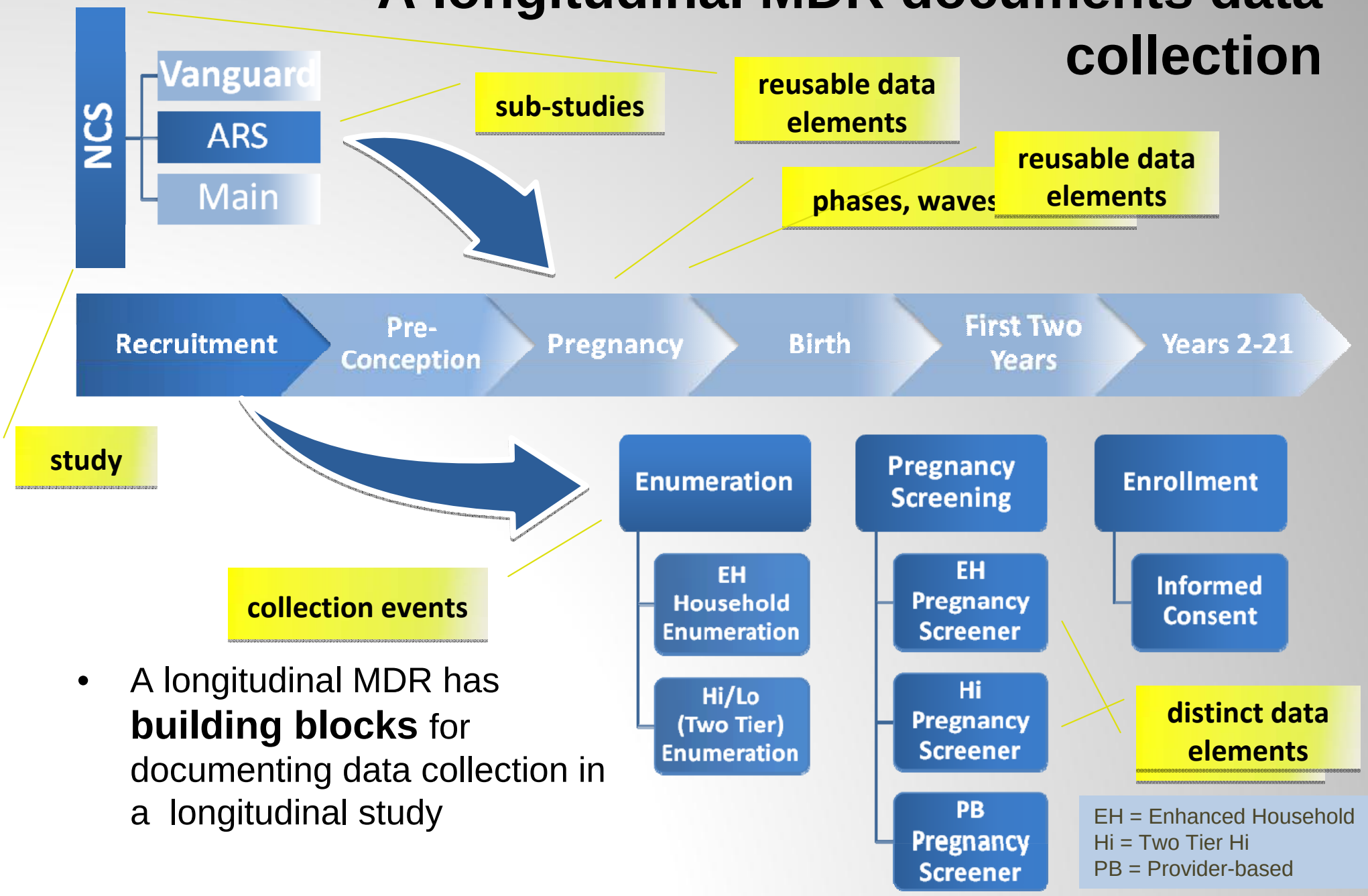
Pre-Conception Follow-up Event Instruments

- PPG CATI
- PPG SAQ

Today's Agenda

- What is longitudinal MDR?
  - A longitudinal MDR documents **data collection** over time
  - A longitudinal MDR documents and facilitates data management over time
- The MDR Website
- Challenges Ahead

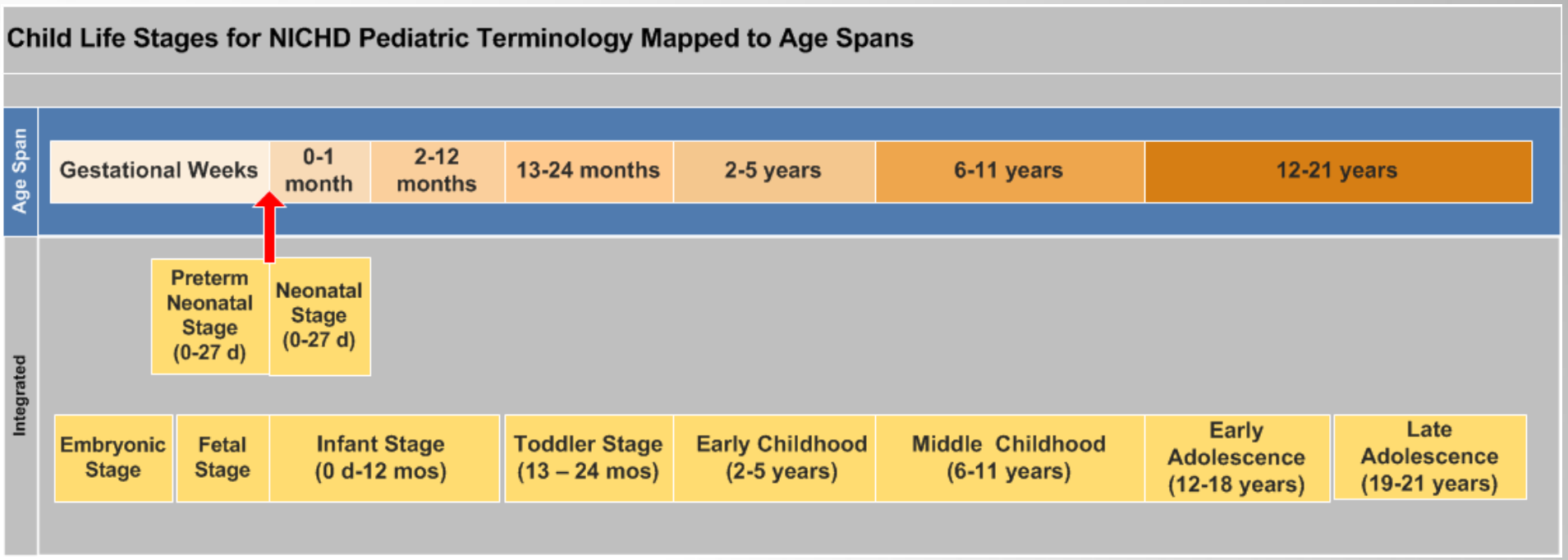
A longitudinal MDR documents data collection



- A longitudinal MDR has **building blocks** for documenting data collection in a longitudinal study

A longitudinal MDR documents data collection (continued)

- In the NCS study phases correspond to child life stages



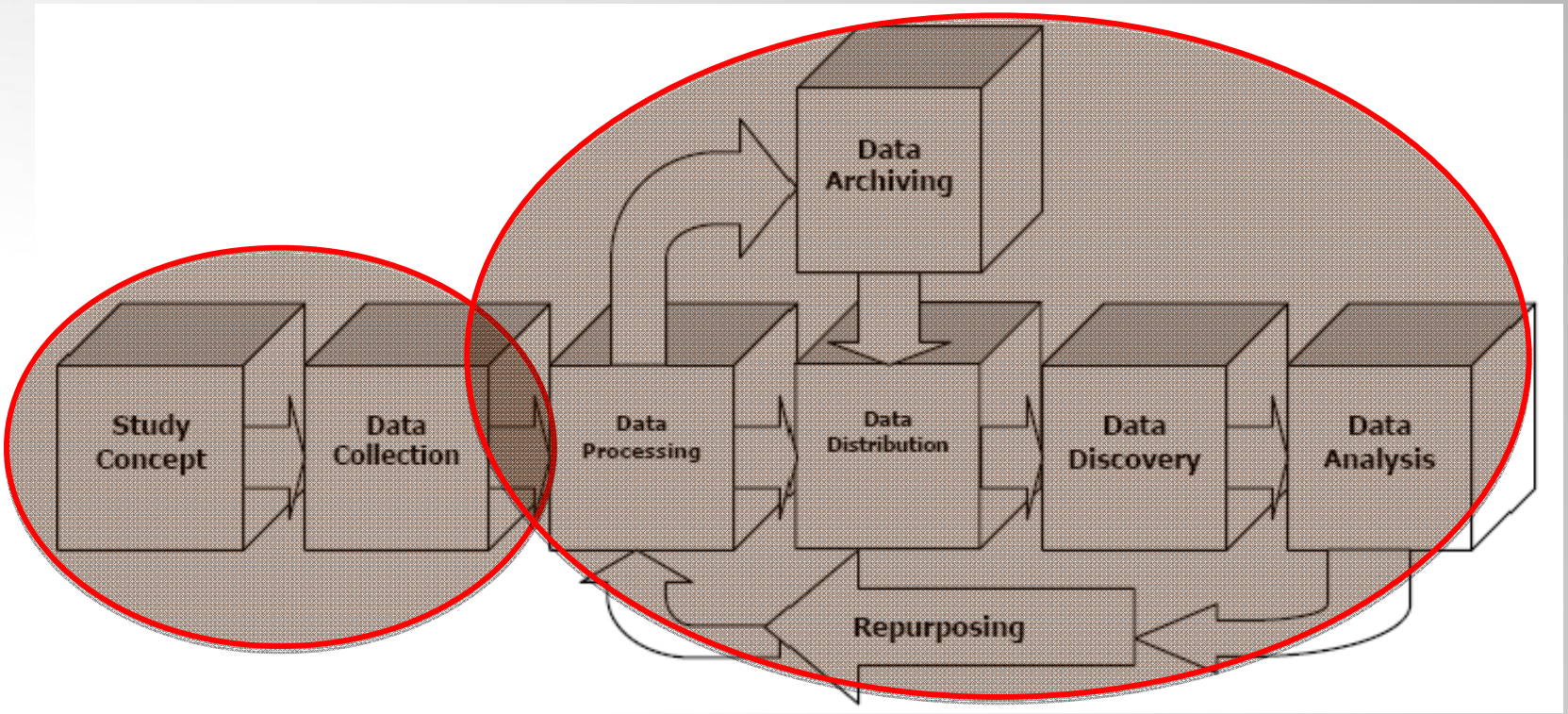
From the NICHD Pediatric Terminology Harmonization Initiative

Today's Agenda

- What is longitudinal MDR?
  - A longitudinal MDR documents data collection over time
  - A longitudinal MDR documents and facilitates **data management** over time
- The MDR Website
- Challenges Ahead

A longitudinal MDR documents data management

- So far we have seen how the MDR describes the study protocol



DDI Data Lifecycle

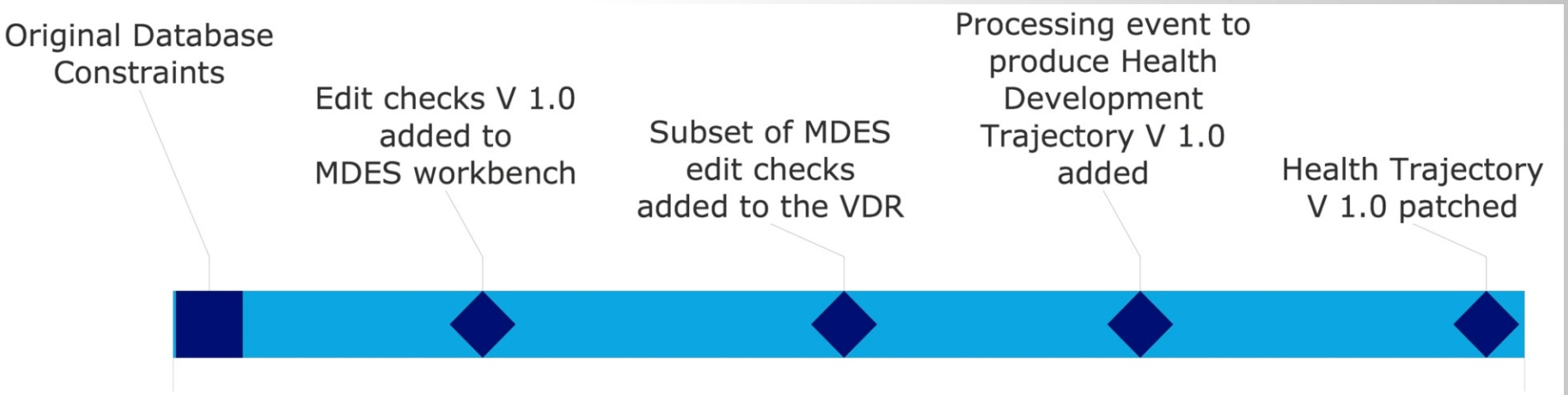
- A study's Concept of Operations (CONOPS) like the NCS CONOPS also specifies **what happens to collected data**
- Note that the "archive" contains a record of the processing events collected data goes through on its way to data distribution, data discovery and data analysis

What are processing events?

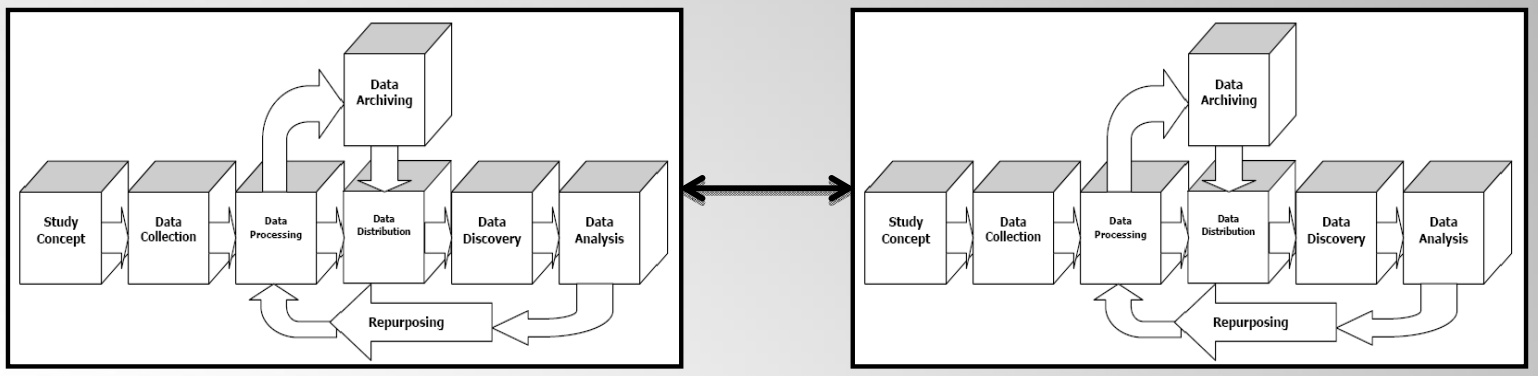
- Edit checks during data collection
- Schema validations
- Edit checks during submission and data integration
- Cleaning operations
- Weighting
- Coding

- Coding in turn includes:
  - Handling non-response questions
  - Recodes
  - Imputation
  - Derivations from multiple questions, variable sources and/or extant data sources

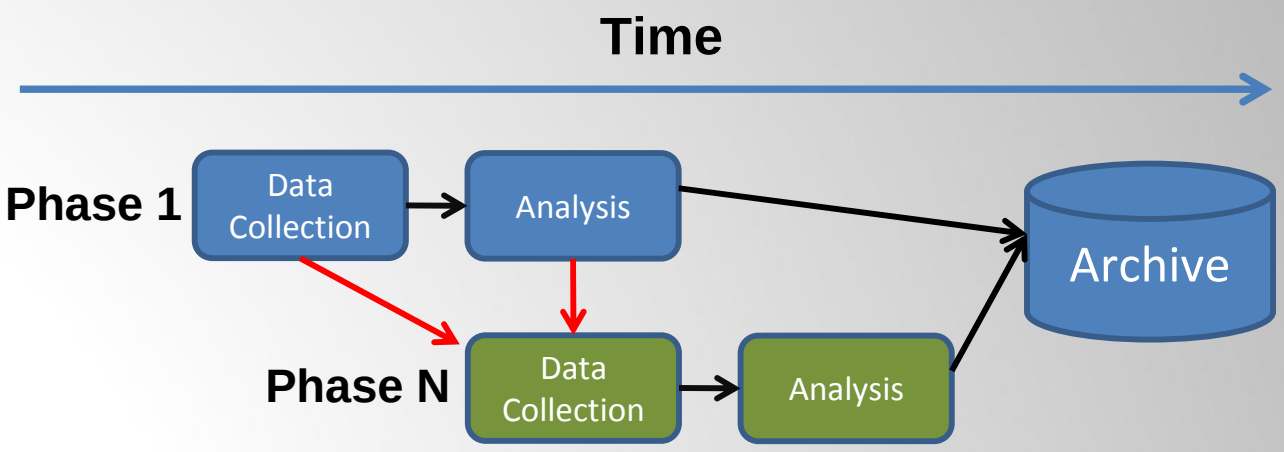
The plan is for the MDR to gather these processing events together in a timeline:



Processing events reach across study phases



- For example:
  - ✓ An analysis in one phase may lead to a change in concepts in the next – e.g. we discover a participant's exposure profile needs more information

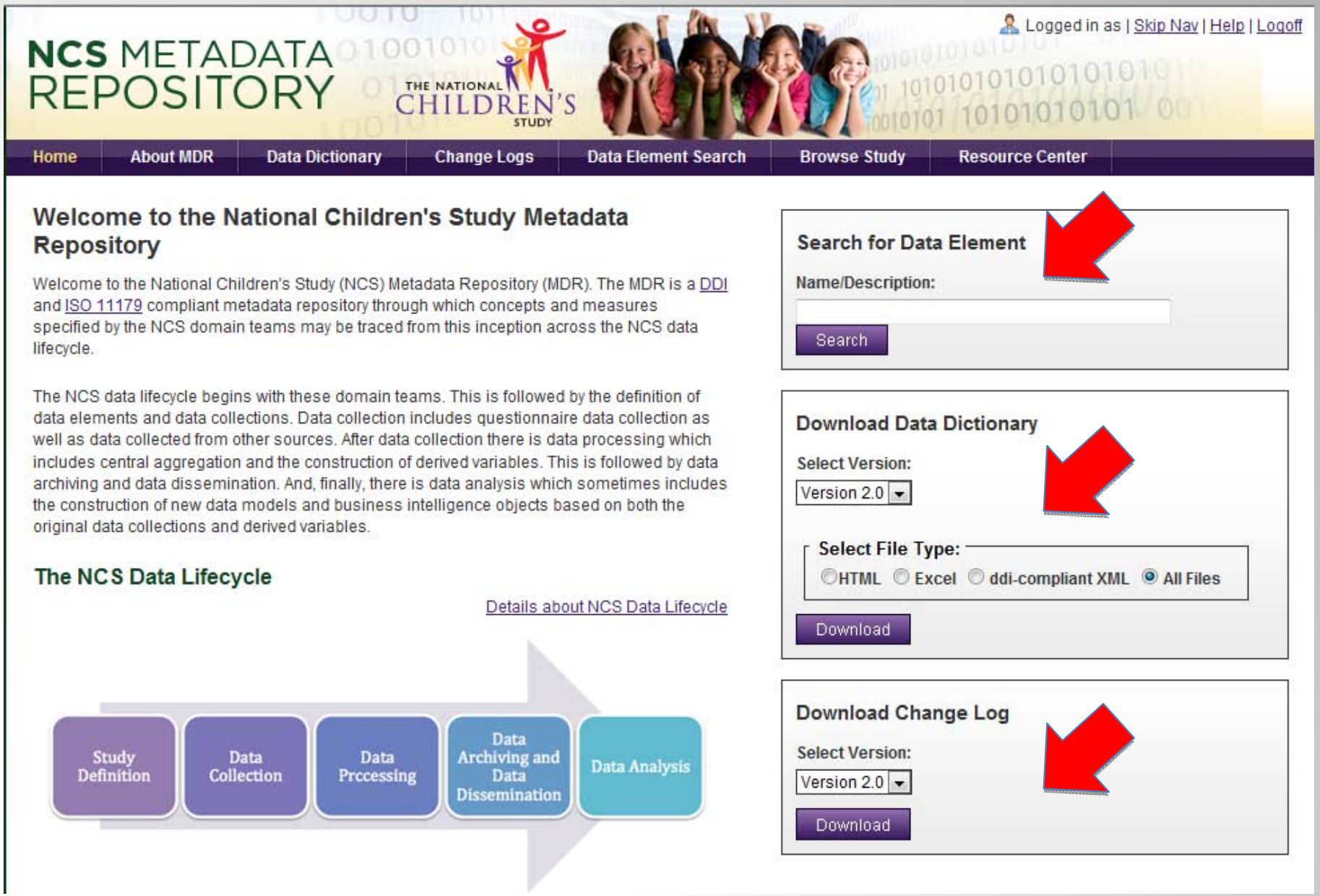


  - ✓ "Core" or, again, repeat measures that correspond to reusable data elements may be "rolled up" for distribution and analysis across study phases
- Again, the archive keeps a record of all these events

Today's Agenda

- What is a longitudinal MDR?
- The MDR Website
  - Screenshots
  - Website Architecture
  - Future RDF Triple Store
- Challenges Ahead

MDR Website Welcome Page



MDR Website Data Dictionary Download Page

NCS METADATA REPOSITORY

THE NATIONAL CHILDREN'S STUDY

Logged in as | Skip Nav | Help | Logoff

Home | About MDR | Data Dictionary | Change Logs | Data Element Search | Browse Study | Resource Center

Data Dictionary

About the DDI Compliant Data Dictionary

The DDI compliant XML data dictionary is an XML representation of a version of the Master Data Elements Specification (MDES) in DDI format. For technical information about this format go to the DDI Specification, where you can download the specification. Note that the data dictionary is in DDI 3.1 format, which is the current version of the DDI specification as of January 2012.
While the DDI specification includes as many as 800 objects, there are elements in the MDES that don't fit into the DDI object model. For example, the position of a data element in a table is not data element attribute in DDI. In cases like this DDI has a note facility. Notes can be attached to each data element. The data dictionary download makes extensive use of the DDI notes facility.

Download Instructions

Each data dictionary is a representation of the Master Data Elements Specification (MDES) that lists the data elements, tables, and code lists. For each version of MDES, a version of the data dictionary is available for download in three formats: **HTML , Excel (which is the MDES), and DDI compliant XML.**

To download a data dictionary, click on the link for the desired version and format(s) in the table below. Note that the downloads will be compressed. The compressed files containing HTML data dictionaries will each contain two files.

Version	HTML (file size)	Excel (file size)	ddi-compliant XML (file size)	All File Types (file size)
2.0	dictionary_2.0_html.zip (1033 KB)	dictionary_2.0_xls.zip (1105 KB)	dictionary_2.0_xml.zip (3727 KB)	dictionary_2.0_all.zip (5809 KB)
1.2	dictionary_1.2_html.zip (490 KB)	dictionary_1.2_xls.zip (658 KB)	dictionary_1.2_xml.zip (1704 KB)	dictionary_1.2_all.zip (2824 KB)
1.1	dictionary_1.1_html.zip (323 KB)	dictionary_1.1_xls.zip (561 KB)	dictionary_1.1_xml.zip (1208 KB)	dictionary_1.1_all.zip (2094 KB)

MDR Website Data Dictionary: HTML Format

NCS METADATA REPOSITORY

THE NATIONAL CHILDREN'S STUDY

Logged in as | Skip Nav | Help | Logoff

Home | About MDR | Data Dictionary | Change Logs | Data Element Search | Browse Study | Resource Center

Data Dictionary

About the DDI Compliant Data Dictionary

The DDI compliant XML data dictionary is an XML representation of a version of the Master Data Elements Specification (MDES) in DDI format. For technical information about this format go to the DDI Specification, where you can download the specification. Note that the data dictionary is in DDI 3.1 format, which is the current version of the DDI specification as of January 2012.

While the DDI specification includes as many as 800 objects, there are elements in the MDES that don't fit into the DDI object model. For example, the position of a data element in a table is not data element attribute in DDI. In cases like this DDI has a note facility. Notes can be attached to each data element. The data dictionary download makes extensive use of the DDI notes facility.

Return to Data Elements Directory
Return to Code List Directory
Return to Table Definitions Directory

Data Element	Table	Data Element Description	Domain Population Concept
B_ADDRESS_1	PREG_VISIT_1	First line of Street Address for Birth Hospital or Birthing Center	PREG_VISIT_1 BIRTHFACILITY STREETADDRESS
BIRTH_PLAN	PREG_VISIT_1	The location where participant plans to deliver baby/babies.	PREG_VISIT_1 PARTICIPANT BIRTHFACILITY

MDR Website Data Dictionary: HTML Format

NCS ME
REPO

Nav | Help | Logoff

Home | Abo

Data Dictio

About the DI

The DDI compliar
about this format
DDI specification
While the DDI sp
element in a table
makes extensive

Download In

Each data diction
version of the dat

To download a da
containing HTML

Version	HTML
2.0	di
1.2	di
1.1	di

information
t version of the

on of a data
ary download

ion of MDES, a

mpressed files

e)
zip (5809 KB)
zip (2824 KB)
zip (2094 KB)

NCIthesaurus

Search ?

Exact Match Begins With Contains

Name/Code Property Relationship

Home | View Hierarchy | Subsets | Help Visited Concepts

Quick Links

Birth Facility (Code C81237)

Suggest changes to this concept

Terms & Properties | Relationships | Synonym Details | View All | View in Hierarchy | View History

Terms and Properties

Definition: The facility which provides a setting with provisions for monitoring and delivering babies.

Preferred Name: Birth Facility

Label: Birth Facility

NCI Thesaurus Code: C81237

NCI Metathesaurus CUI: C2825537 (see NCI Metathesaurus info)

Synonyms & Abbreviations: (see Synonym Details)

Birth Facility

External Source Codes:

UMLS CUI	C2825537 (see NCI Metathesaurus info)

Other Properties:

Legacy_Concept_Name	Birth_Facility
Contributing_Source	NICHD
Semantic_Type	Health Care Related Organization

Additional Concept Data:

Defined Fully by Roles: No

URL: http://ncit.nci.nih.gov/ncitbrowser/ConceptReport.jsp?dictionary=NCI%20Thesaurus&code=C81237

THE NATIONAL CHILDREN'S STUDY
HEALTH GROWTH ENVIRONMENT

llen | Hamilton

MDR Website Data Dictionary: Excel (MDES) Format

Master Data Element Specifications
July 1, 2011
Version 2.0.01

1 = Number
2 = ID
3 = Coded Value

	F	G	H	I	J	K	L
2	PII	Table Order	DE Order	Variable Label/Definition	Type	Numeric Type	For
286		28	18	**External Identifier:** PERSON_ID. Person who withdrew consent. Link to PERSON table. Null if consent has not been withdrawn.	Character		Any string of numbers characters
287		28	19	Status of person who withdrew consent	Numeric	3	1 (Emancipated Minor 2 (Adult at age of majc older) 3 (Parent or Legal Gua
		28	20	Translation method used for consent	Numeric	3	1 (No translation need

Data Elements | Code Lists | Table Description | SC Info | Dispositions | Instrument and Event Map

Normal View | Cell K2 commented by ewissmann | Sum=0

MDR Website Data Dictionary: DDI XML Format


```
960967  <l:Variable id="a89d7d3e-f351-4b7c-8c0a-6e9ad9440b71" version="1.0.0" versionDate="2010-07-01T04:00:00">
960968    <r:UserID type="11179-IRDI">NICHD_Terminology:a89d7d3e-f351-4b7c-8c0a-6e9ad9440b71:1</r:UserID>
960969    <r:UserID type="colectica:sourceId">VS_INSTRUMENT_FATHER_PV1_LIFE_FAILURE</r:UserID>
960970    <l:VariableName xml:lang="en-US">LIFE_FAILURE</l:VariableName>
960971    <r:Description xml:lang="en-US">Father thought his life had been a failure.</r:Description>
960972    <l:ResponseUnit/>
960973    <r:AnalysisUnit/>
960974    <l:Representation>
960975      <l:CodeRepresentation blankIsMissingValue="true">
960976        <r:CodeSchemeReference>
960977          <r:ID>300f79e3-da48-4b98-b28f-582bd064bb13</r:ID>
960978          <r:IdentifyingAgency>NICHD_Terminology</r:IdentifyingAgency>
960979          <r:Version>1</r:Version>
960980        </r:CodeSchemeReference>
960981      </l:CodeRepresentation>
960982    </l:Representation>
960983  </l:Variable>
960984  <l:Variable id="76807f71-0d46-4790-9f6a-b4d1b54a6235" version="1.0.0" versionDate="2010-07-01T04:00:00">
960985    <r:UserID type="11179-IRDI">NICHD_Terminology:76807f71-0d46-4790-9f6a-b4d1b54a6235:1</r:UserID>
960986    <r:UserID type="colectica:sourceId">VS_INSTRUMENT_FATHER_PV1_FELT_FEARFUL</r:UserID>
960987    <l:VariableName xml:lang="en-US">FELT_FEARFUL</l:VariableName>
960988    <r:Description xml:lang="en-US">Father felt fearful.</r:Description>
960989    <l:ResponseUnit/>
960990    <r:AnalysisUnit/>
960991    <l:Representation>
960992      <l:CodeRepresentation blankIsMissingValue="true">
960993        <r:CodeSchemeReference>
960994          <r:ID>300f79e3-da48-4b98-b28f-582bd064bb13</r:ID>
960995          <r:IdentifyingAgency>NICHD_Terminology</r:IdentifyingAgency>
960996          <r:Version>1</r:Version>
960997        </r:CodeSchemeReference>
```

MDR Website Browse Study Page

NCS METADATA REPOSITORY

THE NATIONAL CHILDREN'S STUDY

Logged in as Art Vandelay | Skip Nav | Help | Logoff

Home | About MDR | Data Dictionary | Change Logs | Data Element Search | Browse Study | Resource Center

Browse Study

Operational Data Elements | Instrument Data Element

Recruitment | Pre-Conception | Pregnancy | Birth | First Two years | Years 2-21

Pre-Conception is a series of collection events. During this period study participants who were not already pregnant at the time of enrollment are followed to determine if they have become pregnant. Initially the participant's Pregnancy Probability Group (PPG) is calculated. Follow-ups are scheduled based on the participant's PPG status.

Each phase of the study has one or several collection events at which data is collected. A collection event could be an interview, phone call, or a visit in which an instrument(s) is administered. For each collection event, there is a collection of instruments that are either administered at the same time or are more than one version of an instrument. Collection events and collections of instruments that are administered at each collection event are listed below. To view a list of tables within a collection of instruments, click on the link for the desired collection of instruments below.

Collection Events

Pre-Conception Visit Event Instruments

- Pre-Pregnancy Visit

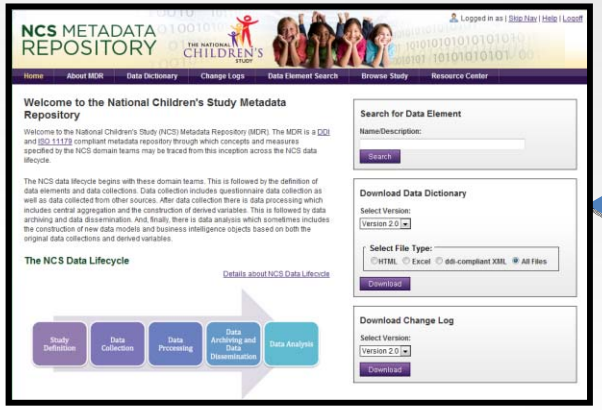
Pre-Conception Follow-up Event Instruments

- PPG CATI
- PPG SAQ

Today's Agenda

- What is a longitudinal MDR?
- The MDR Website
  - Screenshots
  - Website Architecture
  - Future RDF Triple Store
- Challenges Ahead

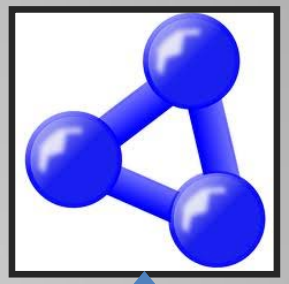
The MDR website presents the study framework to study planners and researchers



MDR Website Database

In the future this store can support reasoning over study objects using SPARQL queries

RDF Triple Store



Concepts and Instructions

Master Data Elements Specification

Longitudinal MDR Database

Operational Data Elements

Instrument Data Elements

Processing Events

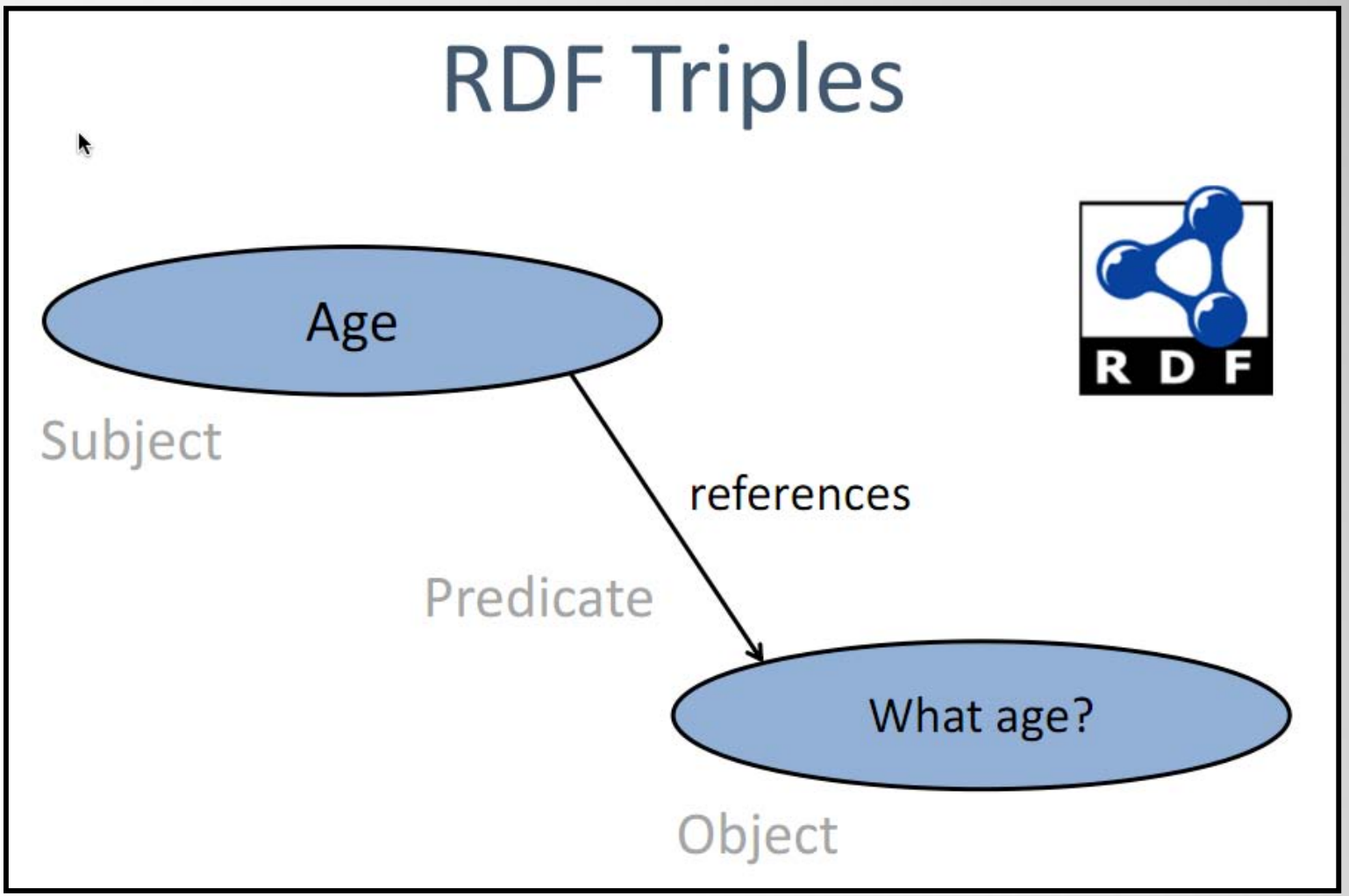
Extant Data Definitions

Today's Agenda

- What is longitudinal MDR?
- The MDR Website
  - Screenshots
  - Website Architecture
  - Future RDF Triple Store
- Challenges Ahead

Future RDF Triple Store

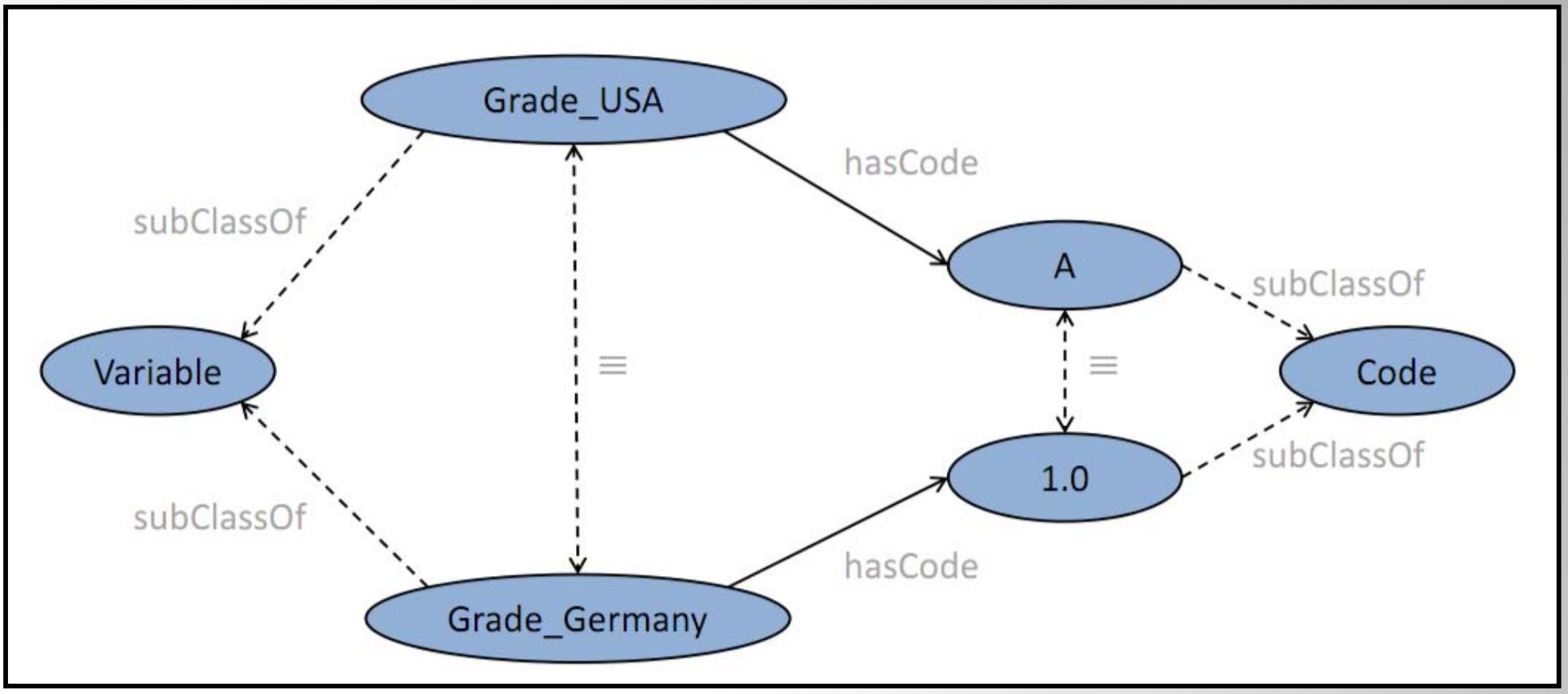
- What is the Resource Description Framework (RDF)?



From Representation of the Data Documentation Initiative Using Semantic Web Technologies
by Thomas Bosch

Future RDF Triple Store (continued)

- What kinds of problems can be solved?



From Representation of the Data Documentation Initiative Using Semantic Web Technologies by Thomas Bosch

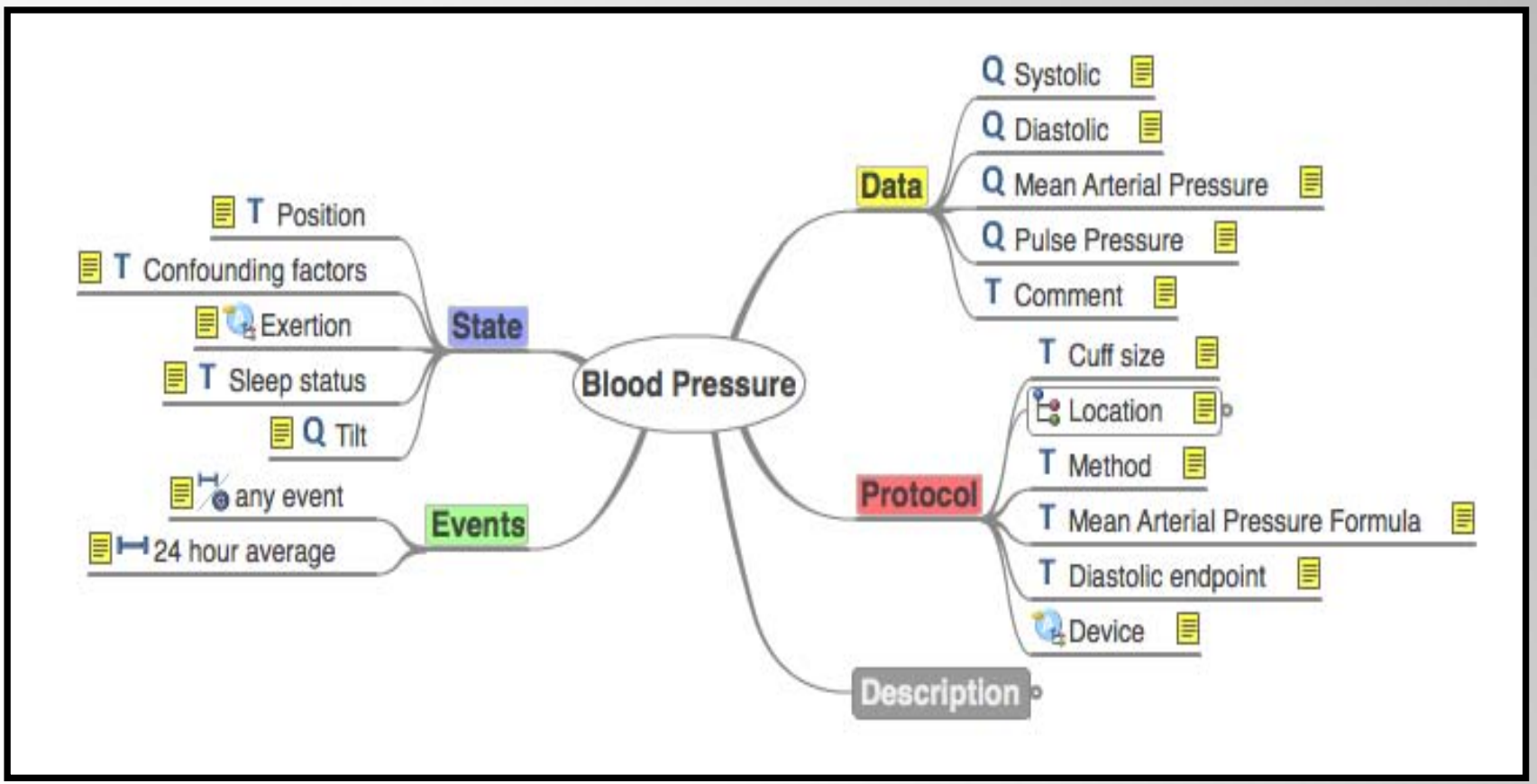
The store can be queried to explore the comparability of coded data elements

Today's Agenda

- What is longitudinal MDR?
- The MDR Website
- Challenges Ahead
  - Identifying and Maintaining Reusable Data Elements
  - Representing Other Data Collection Besides Questionnaires
  - Building, Maintaining and Using Participant Profiles

Identifying and Maintaining Reusable Data Elements

1. Identifying and maintaining reusable data elements across different child life stages and over time as technologies, devices and protocols change…



From openEHR Clinical Manager

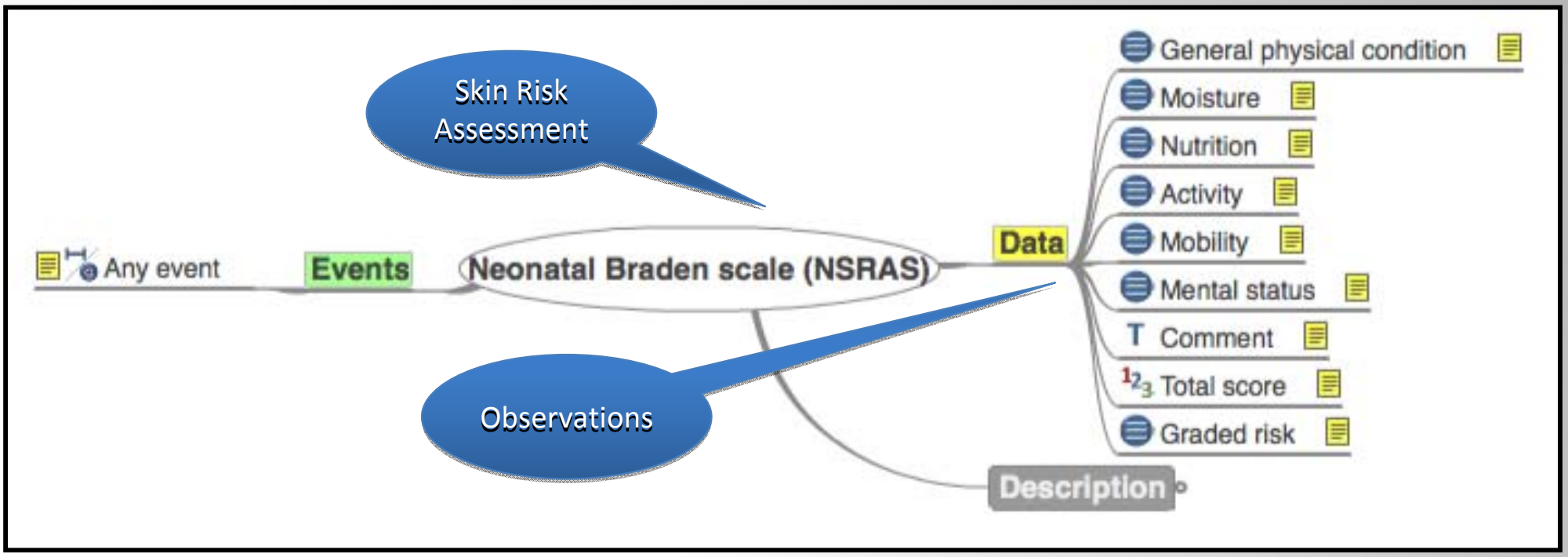
*By extending our descriptions of data elements to include not just the **what** but also the **how** and the **when** we are better positioned to evaluate comparability over time*

Today's Agenda

- What is longitudinal MDR?
- The MDR Website
- Challenges Ahead
  - Identifying and Maintaining Reusable Data Elements
  - Representing Other Data Collection Besides Questionnaires
  - Building, Maintaining and Using Participant Profiles

Representing Other Data Collection Besides Questionnaires

2. Representing other data types besides questionnaires in the data collection stage of operations…



From openEHR Clinical Manager

In standardized tests and exams answers are often weighted and aggregated. Weightings can change over time as tests and exams are re-standardized

Representing Other Data Collection Besides Questionnaires

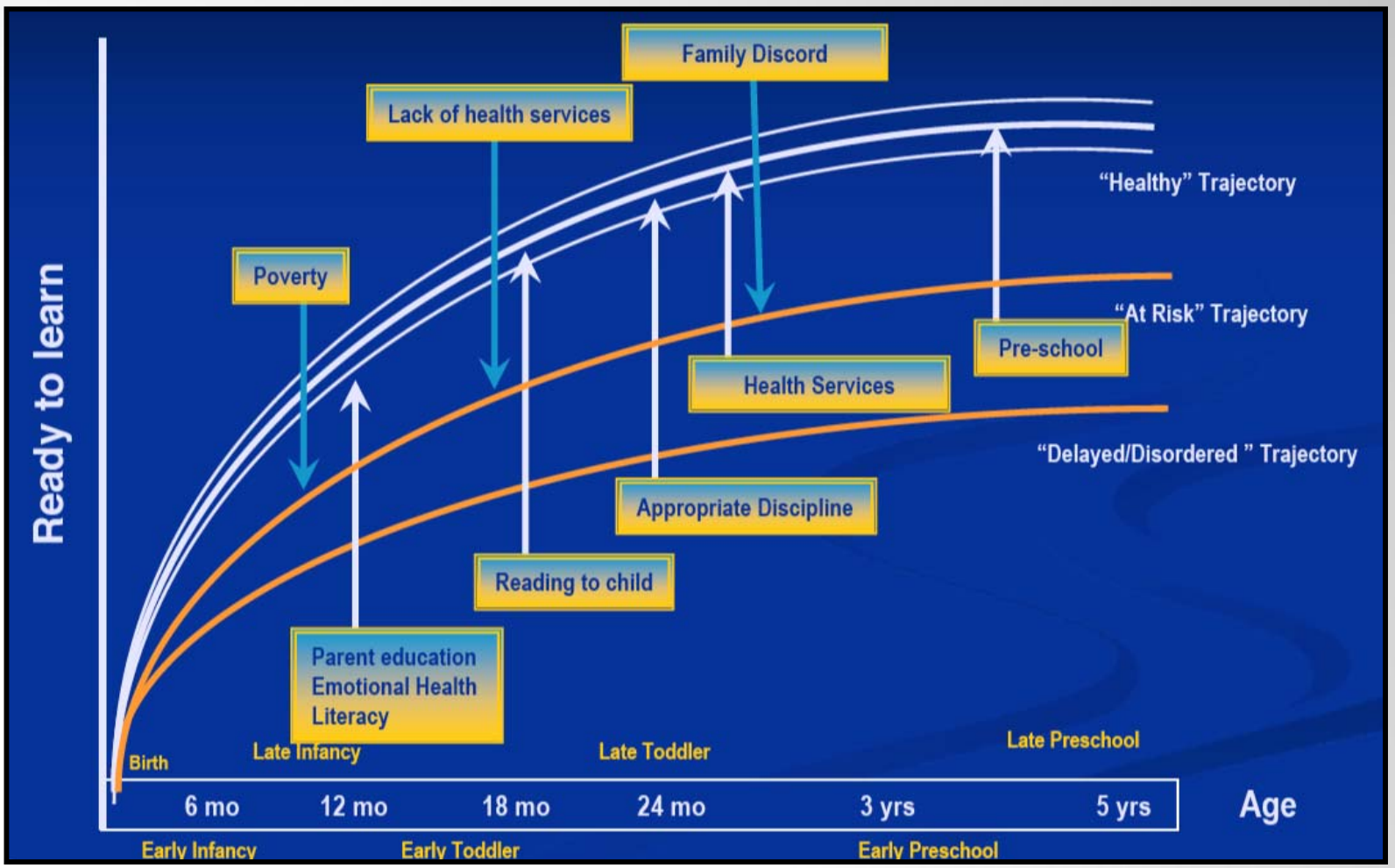
		significant changes independently.] 4: Completely immobile [Does not make even slight changes in body or extremity position without assistance (e.g., Pavulon).]
Mental status Ordinal Occurrences: 0..1 (optional)	An evaluation of the mental status or sensory preception of the neonate.	1: No impairment [Alert and active.] 2: Slightly limited [Lethargic.] 3: Very limited [Responds only to painful stimuli (flinches, grasps, moans, increased blood pressure or heart rate).] 4: Completely limited [Unresponsive (does not flinch, grasp, moan, increase blood pressure, or heart rate) to painful stimuli due to diminished level of consciousness or sedation.]
Comment Text Occurrences: 0..1 (optional)	An additional comment about the neonatal pressure risk assessment.	Free or coded text
Total score Count Occurrences: 0..1 (optional)	The total score, derived from the sum of all sub-scores for identifying if a neonate is at risk for pressure ulcers.	min: >=2; max: <=24
Graded risk Ordinal Occurrences: 0..1 (optional)	A graded risk derived from the summed Total score.	0: Not at risk - Total score less than 13 [The neonate is at risk of developing a pressure ulcer.] 1: At risk - Total score 13 or over [The neonate is not at risk of developing a pressure ulcer.]

Today's Agenda

- What is longitudinal MDR?
- The MDR Website
- Challenges Ahead
  - Identifying and Maintaining Reusable Data Elements
  - Representing Other Data Collection Besides Questionnaires
  - Using Participant Profiles Over Time

Using Participant Profiles Over Time

3. Objects like health development trajectories may be used in an adaptive protocol to trigger prospective studies for subsets of participants



How can we be sure that as we grow these objects based on new knowledge that our prospective studies don't contain apples and oranges?

From Conceptual Models to Guide Research by Neal Halfon (NCS Health Measurements Network)

Next Steps

- Continue to develop the infrastructure of Longitudinal MDR that will be needed regardless of specific capabilities
- Using inputs from this workshop, work with Dr. Hirschfeld, Dr. Haugen, Social and Scientific Systems and others to prioritize and plan future capabilities and releases