CLINICAL DATA INTERCHANGE STANDARDS CONSORTIUM

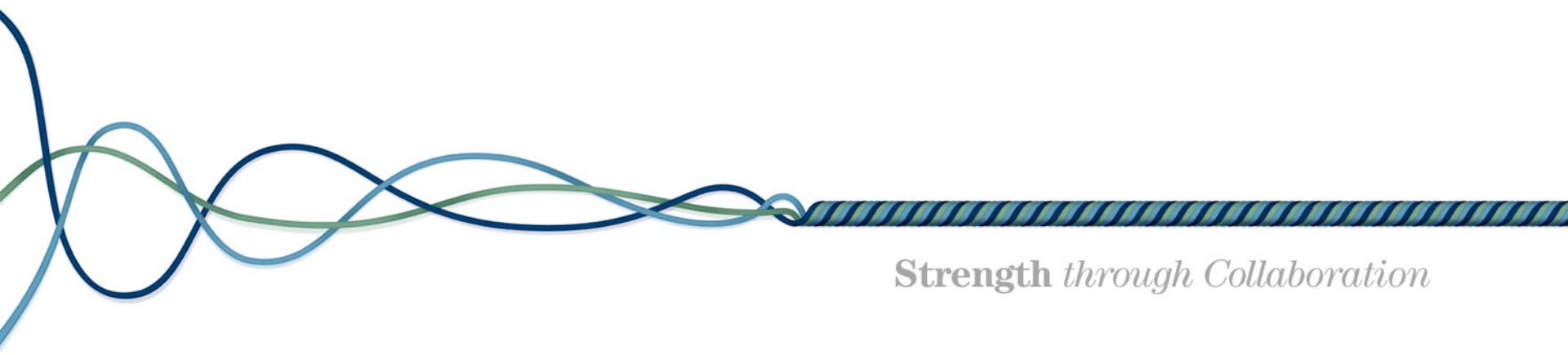
The CDISC vision is to inform patient care & safety through higher quality medical research.

Strength *through Collaboration*

Null Flavors

A tool for handling missing and awkward data

Presented by Diane Wold



Strength *through Collaboration*

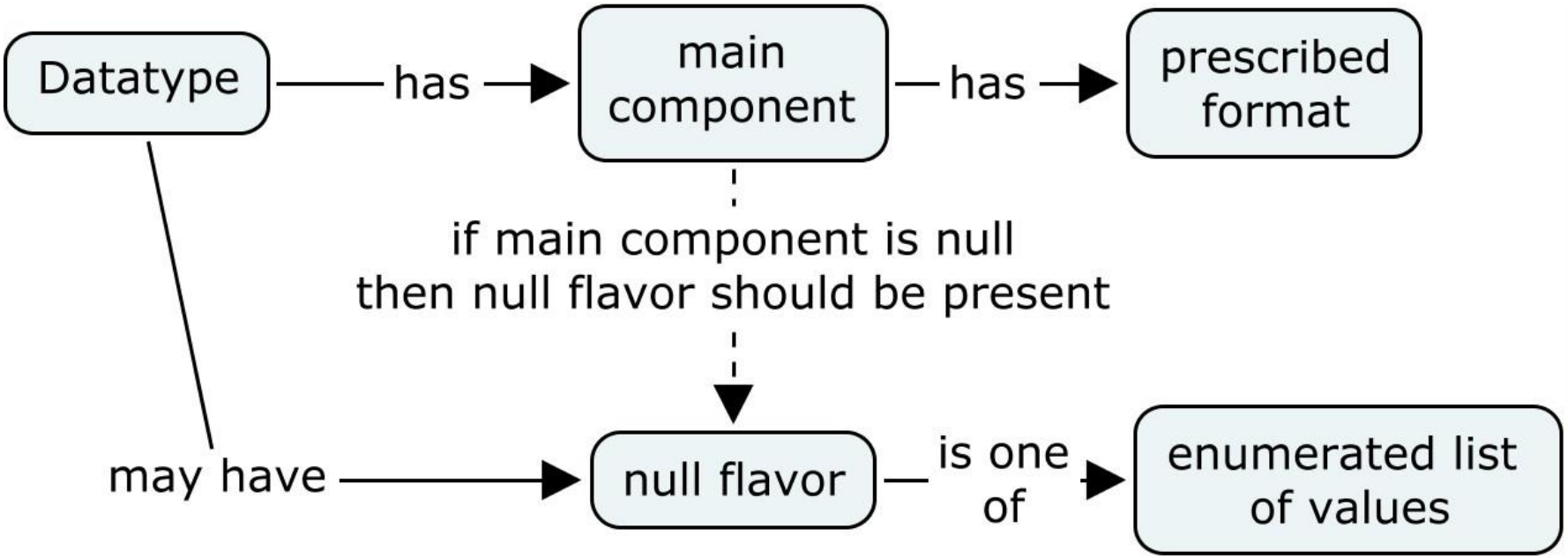
Outline

- Source – ISO and HL7 Data Type Standard
- Null Flavor Hierarchy
- “Unknown” Null Flavors
- “Invalid” Null Flavors
- Null Flavors in the SDTM world
- Future Use of Null Flavors
- Why Null Flavors Now?

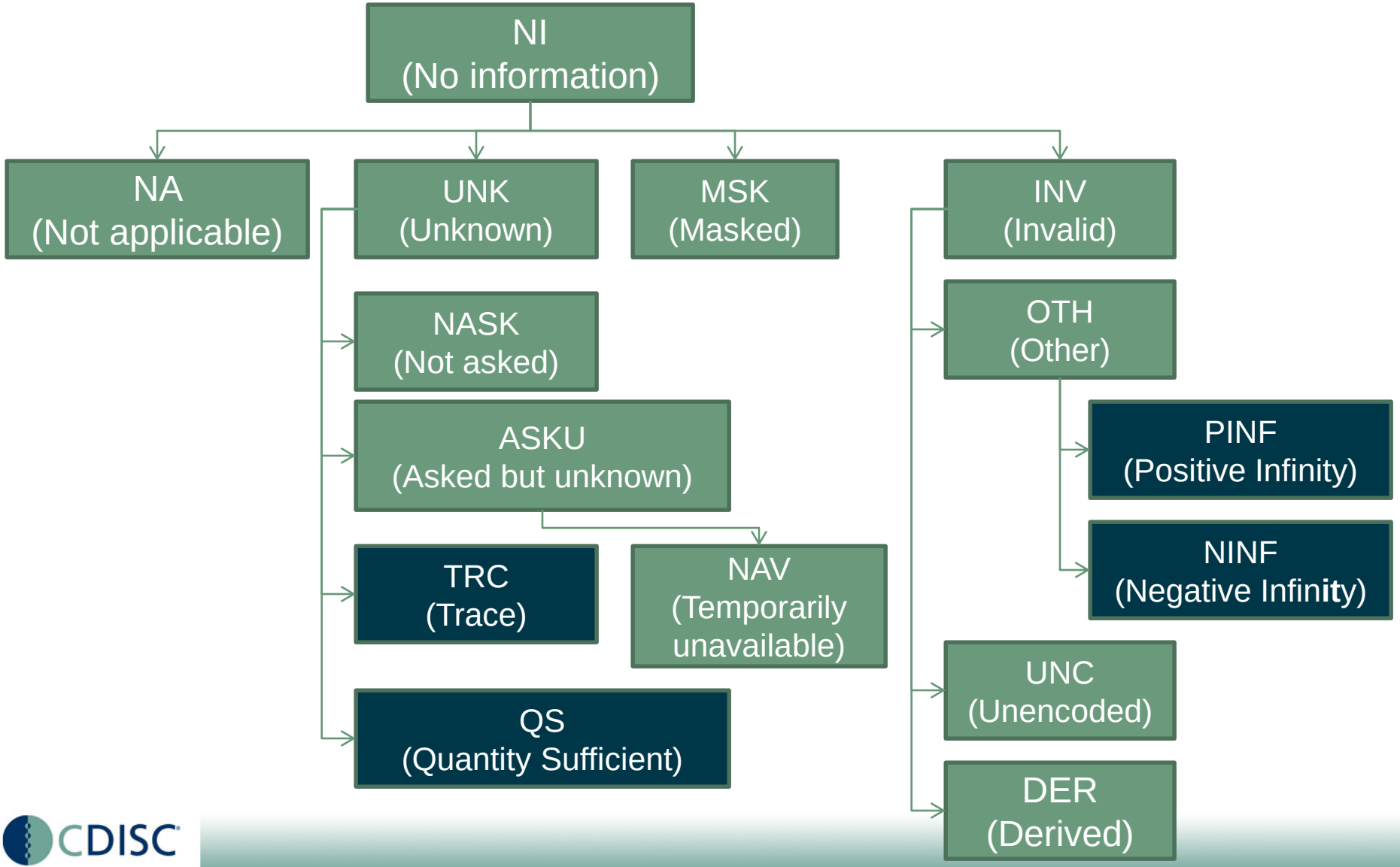
Complex Data Types

- Data Types for Health Care
  - ISO 21090
  - HL7 R2 Data Types
- Complex data types have multiple components
  - Null flavor is a component that accompanies the “main” component of a data type.
  - The null flavor is populated when the main component is null, and provides a little information about that null main component
- Data type standard is expressed in xml
  - Easily handles as many components are as needed for any particular situation

Null Flavor Component of a Complex Data Type



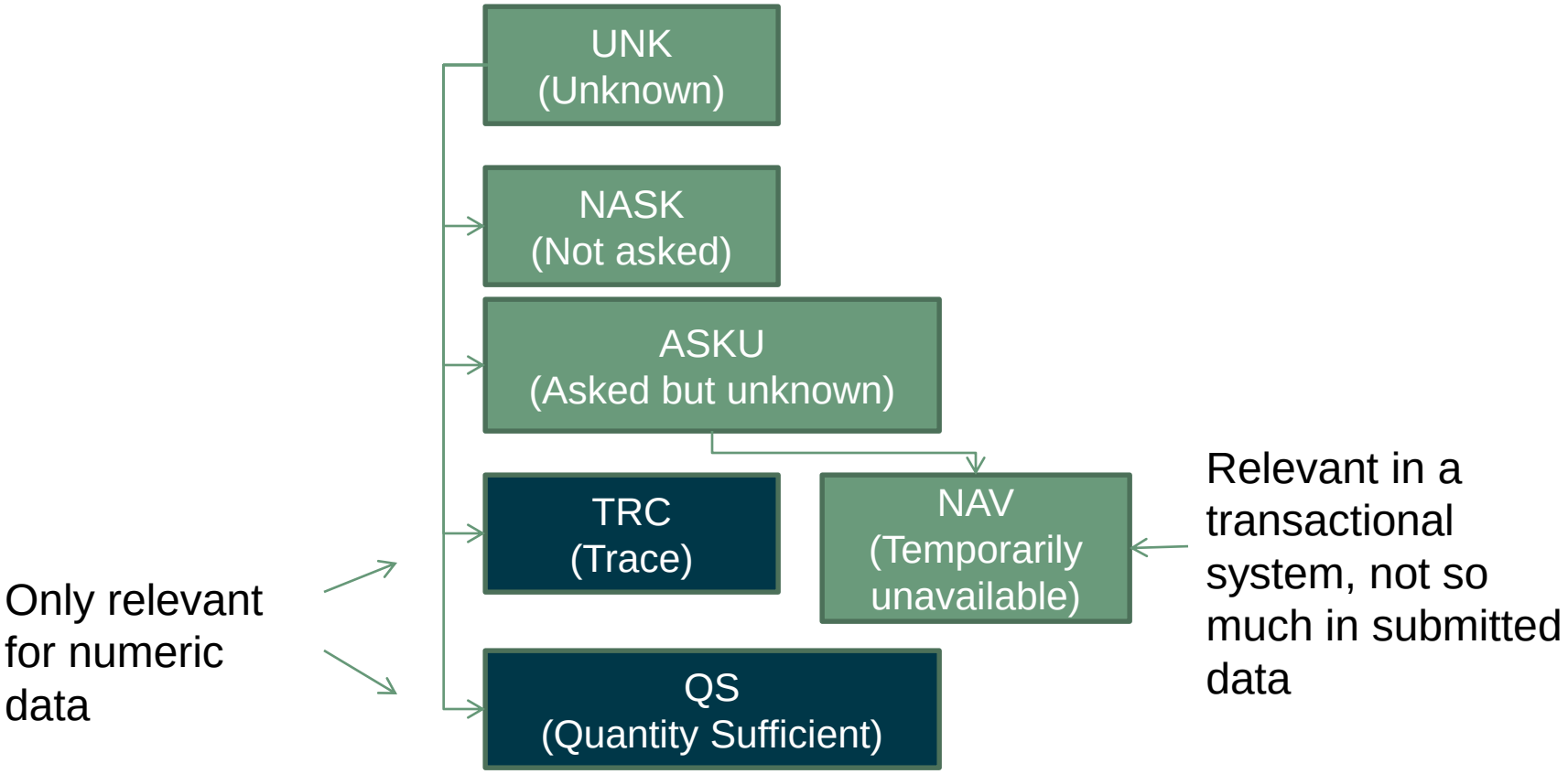
Null Flavor Hierarchy



Use cases for Null Flavors NI, NA, and MSK

- NI (No information)
  - No input to a field in a form
  - Data out of scope of a system
    - If ethnicity was not collected for a particular trial, then the value for the SDTM variable ETHNIC would be null, and the associated null flavor would be NI
- NA (Not applicable)
  - Menopausal status for a male subject
  - Home address for a homeless person
- MSK (Masked)
  - Subject treatment assignment before unblinding
    - If interim data on a study was submitted before unblinding, the SDTM variable ARM would be null, the associated null flavor would be MSK

“Unknown” Null Flavors



Use Cases for “Unknown” Null Flavors

- UNK (Unknown)
  - Subject is female, so menopausal status is applicable, but is unknown.
- NASK (Not asked)
  - Patient Reported Outcome was not administered
- ASKU (Asked but unknown)
  - Patient Reported Outcome was administered, but some questions were not answered
  - Subject was asked about history of chicken pox, but didn’t know whether they had it
- NAV (Temporarily Unavailable)
  - Subject forgot to bring remaining study medication to a clinic visit, but is expected to bring it to the next visit

Queries about blank fields, a possible use for Unknown null flavors

- For CRF fields that are thought to be important enough, queries are commonly issued when the field is left blank
- If sites can pro-actively record a null flavor, these queries would be avoided
- In fact, this is why null flavors like “Unknown” and “Not applicable” are often included in lists of responses, even though these are not real answers
  - The more detailed NASK, ASKU, or NAV might provide additional value beyond UNK
  - When the blank field is the result of a finding, the SDTM variable REASND provides a place for more detail about the null result, could be populated with these null flavors

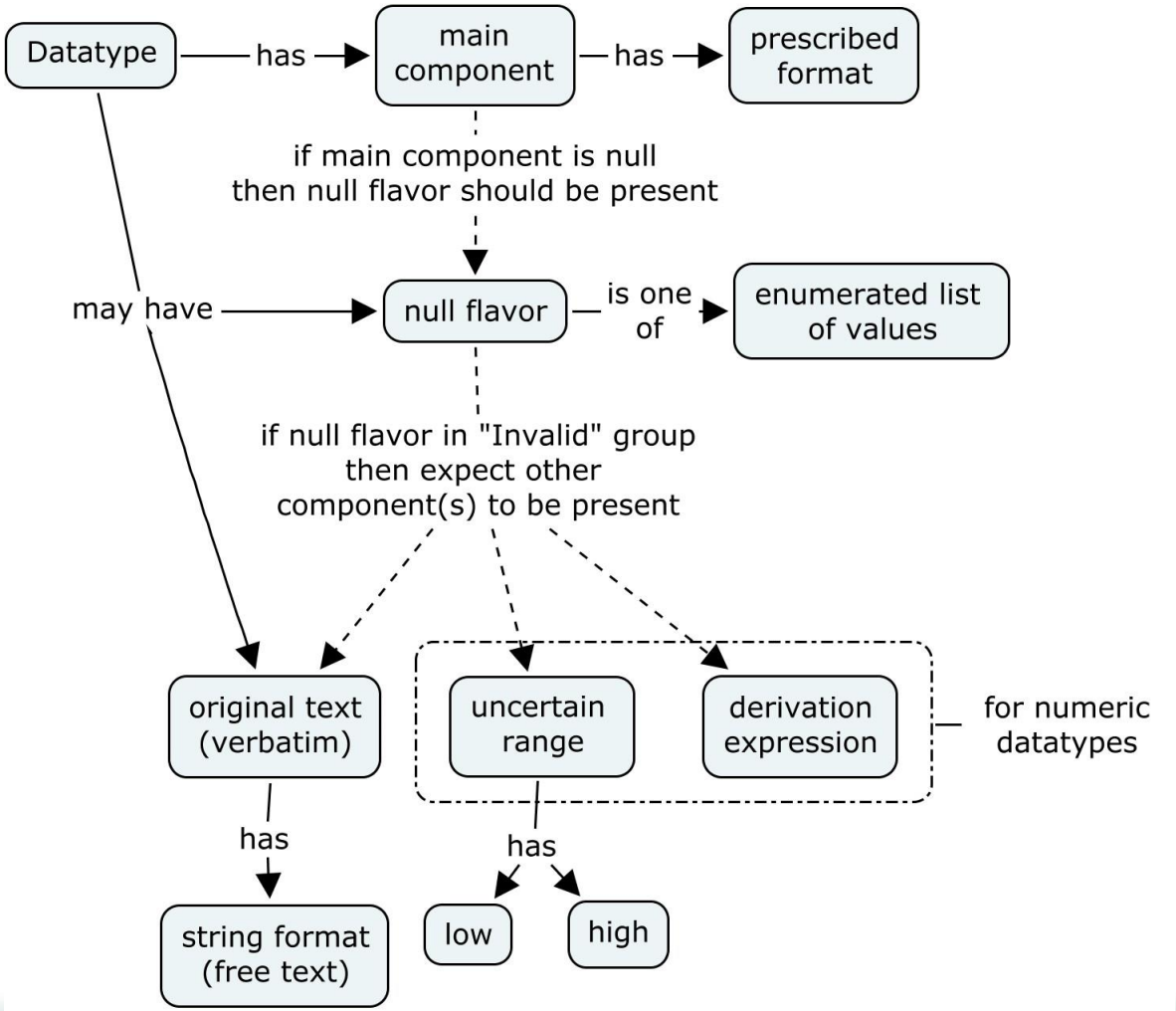
“Trace” and “Quantity Sufficient”

Trace	The content is greater than zero, but too small to be quantified
Quantity Sufficient	The specific quantity is not known, but is known to be non-zero and is not specified because it makes up the bulk of the material. “Add 10 mg of ingredient X, 50 mg of ingredient Y, and sufficient quantity of water to 100 ml.” The null flavor would be used to express the quantity of water.

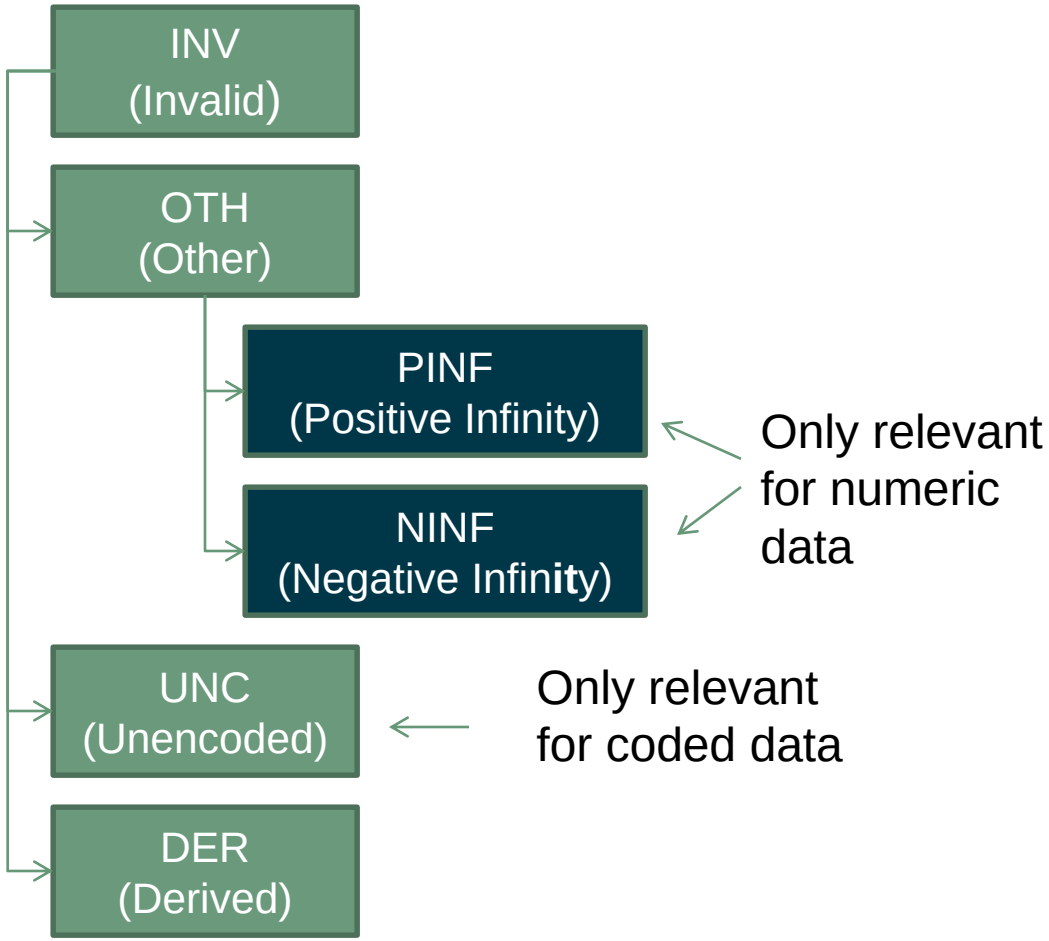
Use case for TRC

- SDTMIG 4.1.5.1.3, lab example, row 6
- Ketones result is given as BLQ
  - In xml: <value nullFlavor =“TRC” />
  - Or: <value nullFlavor =“TRC” originalText=“BLQ”/>
- SDTM might provide the lower limit of quantitation in the SDTM variable LBLLOQ, which provides additional information about the meaning of the null flavor

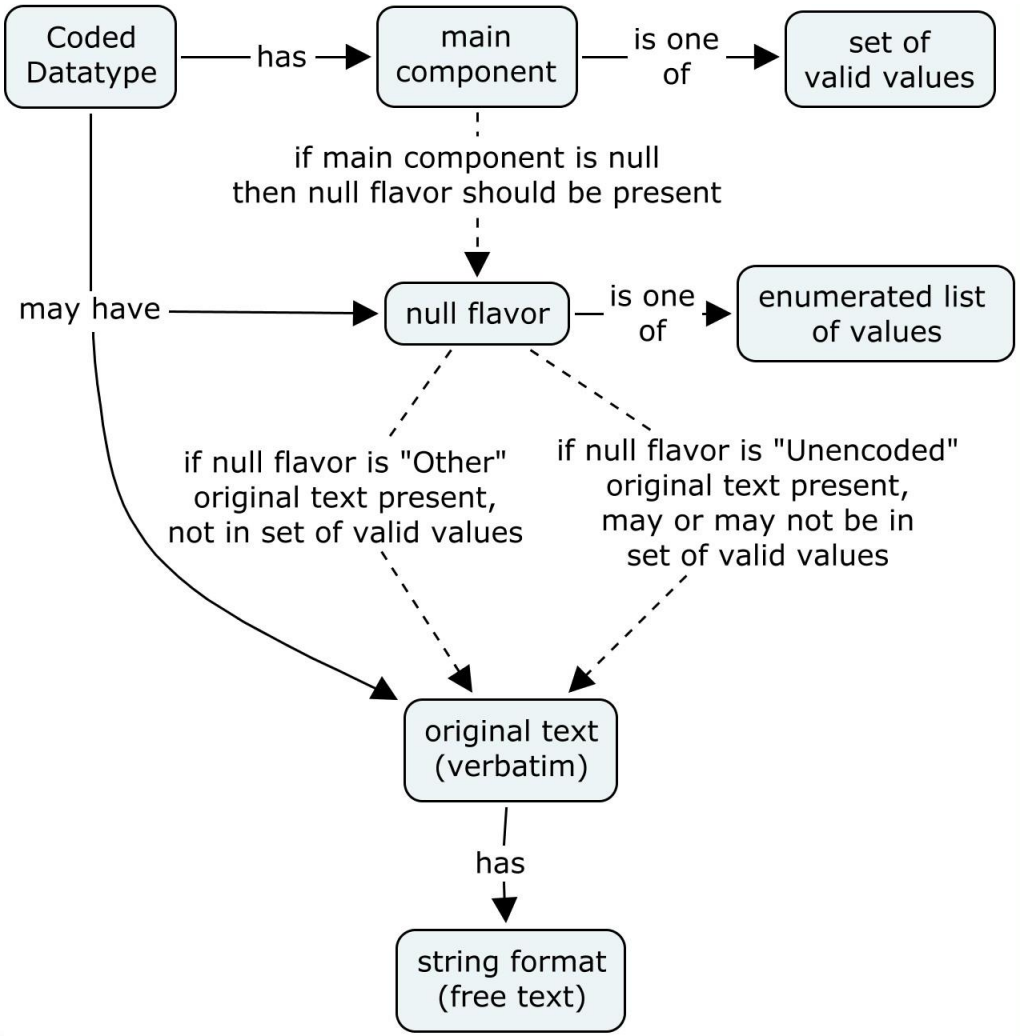
Datatype Components Which May Contain "Invalid" Data



“Invalid” Null Flavors



Null Flavors and Original Text for Coded Data



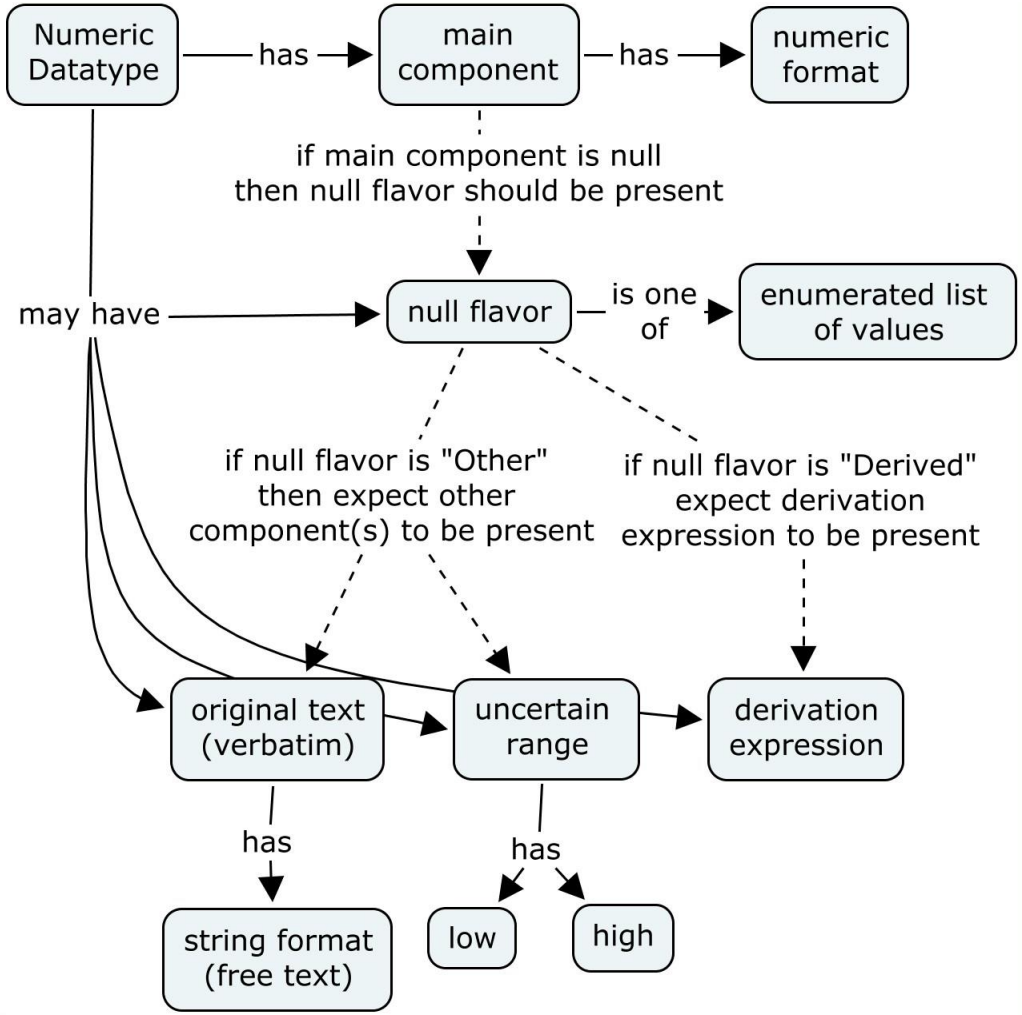
Use cases for “Other” Null Flavor in Coded Data

- A response other than one of the valid values in a code list has been selected
  - SDTMIG Section 4.1.2.7.1 example: CRF has four choices for injection location (Right arm, Left arm, Right thigh, Left thigh), but injection was administered in upper right abdomen
  - In xml:

    <value nullFlavor=“OTH”>
    <originalText=“upper right abdomen”/>
    </value>
  - In this case, we have a null flavor and an original text, but only one SDTM variable (EXLOC). The SDTMIG provides three choices about how to handle the null flavor (Other) and the original text (specify text).
- Concomitant medication data was not coded
  - In xml:

    <value nullFlavor=UNC>
    <originalText=“aspirin”/>
    </value>
  - SDTM variable CMDECOD is null, the associated null flavor is UNC. The SDTM variable CMTERM contains the original text (verbatim).
  - SDTM is prepared for both original text and the decode for the topic variable in CM, but doesn’t have a place for the UNC null flavor.

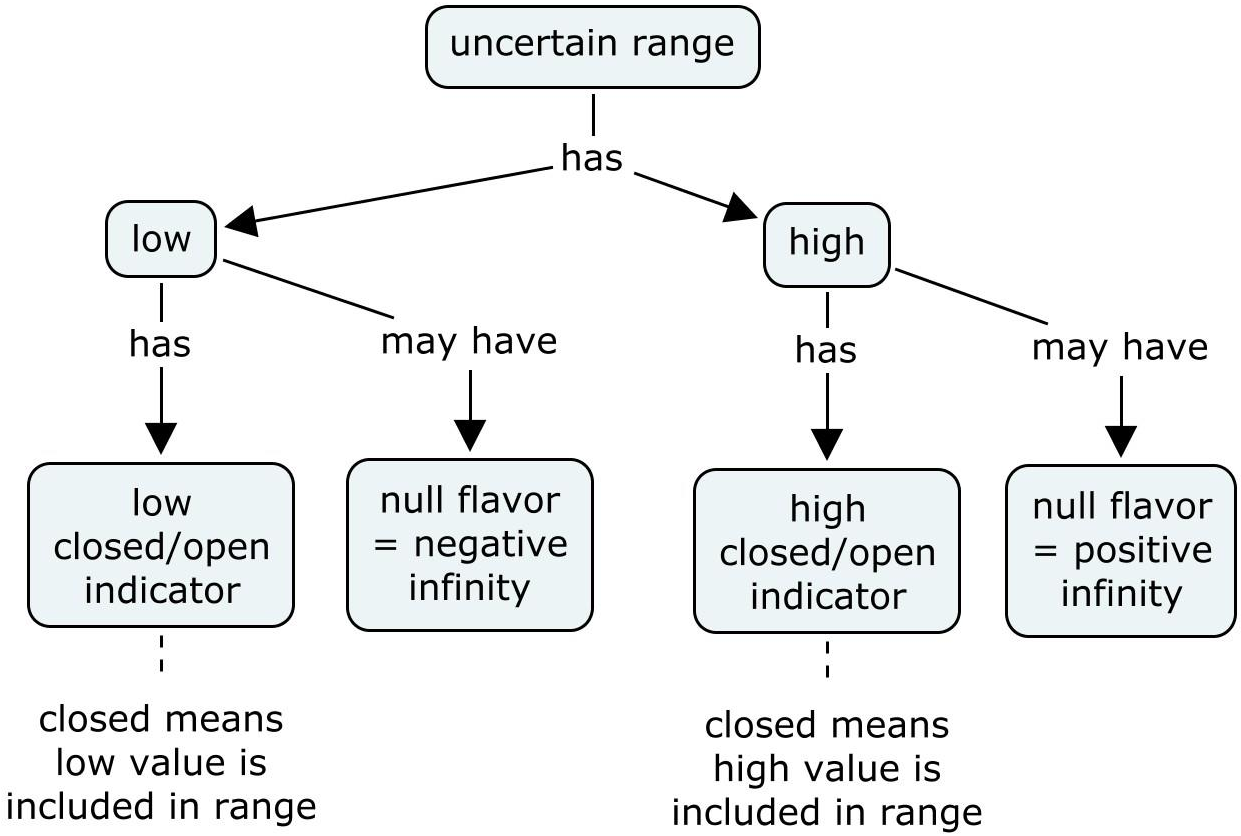
Null Flavors and Datatype Components for Numeric Data



Use cases for Other Null Flavors in Numeric Data

- WBC result is given as “Abnormal high”
  - In xml: <value nullFlavor=“OTH” originalText=“Abnormal high”/>
  - SDTM captures the original text in LBORRES, leaves LBSTRESN null
- HEIGHT = 175 cm, WEIGHT = 80 kg. BMI is null
  - In xml: <value nullFlavor=“DER” expression=“WEIGHT/(HEIGHT^2)”/>

Uncertain Range and "Infinity" Null Flavors



Use cases for Other Null Flavors in Numeric Data (2)

- WBC result is given as “3000-4000”
  - In xml:

```
<value nullFlavor = “OTH”
    <uncertainRange>
      <low  value=“3000”/>
      <high value=“4000”/>
    </uncertainRange>
/>
```

- SDTMIG 4.1.5.1.3 lab example, row 11: WBC result given as “<4,000”
  - In xml:

```
<value null flavor =“OTH” originalText=“<4,000”
     <uncertainRange>
       <low nullFlavor =“NINF”>
       <high value=“4,000”/>
       <highClosed=“FALSE”/>
     </uncertainRange>
/>
```

  - If it is recognized that WBC is never negative, <low nullFlavor=“NINF”> could be replaced by <low value=“0”>

Null Flavors in SDTM

- In a few places, SDTM separates data type components
  - Numeric value & unit
  - Verbatim, code, and decode
- In most places, null flavors are placed in the same field as “proper” values
  - --ORRES is character to handle all data types, all formats
  - Code lists often include null flavors (Other, Not applicable, Unknown) along with “proper” values

Null Flavors in SDTM (2)

- --STAT is a pseudo-null flavor
  - The only valid value, “NOT DONE,” is used to cover more than the closest null flavor, “Not asked”
- Null flavors are implemented for the Trial Summary domain
- SDTMIG provides a little, scattered advice, on things like values of “other” and associated “specify” text

Future use of Null Flavors

- In tabular representations of data (like SDTM and ADaM), separating null flavors from “proper” values requires adding more columns
  - Null flavors are just one component of the complex data types, accommodating other components would mean even more columns
  - The complex data types have nested components, as we saw with the uncertainRange
- Full implementation of complex data types in tabular data is not feasible
- Using the ideas in the complex data types can add value
  - More consistent conventions for handling null flavors, original text, uncertain ranges, etc.
  - Example: SAS missing value codes are a mechanism somewhat similar to null flavors

Future use of Null Flavors (2)

- Other representations (like xml and RDF) can include relevant components as needed
- As CDISC moves toward these technologies, richer data types can be implemented
  - We already see that ODM and define.xml can store richer data type information than the Num and Char of SAS datasets

Why Null Flavors Now?

- Provide a framework for thinking about awkward and missing data
- Null flavor terms provide controlled terminology that can be used in codelists, including sponsor controlled terminology
  - CDISC CT currently includes either “NOT APPLICABLE” and “NA” in codelists
- Provide an introduction to complex datatypes, which provide a rich framework for understanding data types and how they are handled in current and future standards