Telling the Data Story: Use of Informatics, Harmonized Semantics and Metadata in the National Children’s Study

John Lumpkin, MS, MBA, PMP
Steven Hirschfeld, MD, PhD

NIH-NICHD-National Children’s Study Program Office

October 19, 2012
What are some of the gaps?

- Health Care System || Clinical Research
- Individual Research Study || Research Study
- People || Technology
- Convention || Precision
- Custom Solutions || Standards-based Solutions
- Present || Future
The National Children’s Study (NCS) is mandated by the U.S. Congress and implemented by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) with input from a consortium of federal partners.

The NCS is a multi-year study that will examine the effects of environmental influences on the health and development of more than 100,000 children across the United States from before birth until 21 years of age.

The goal of the NCS is to improve the health and well-being of children and contribute to the understanding of how various factors impact health and disease.
The NCS is an integrated system of activities that include:

- a Pilot Study which began in January 2009 with the goal of determining the feasibility, acceptability and cost of Study activities,
- a Main Study scheduled to begin in 2012 to investigate exposure-response relationships and
- various sub-studies and formative research projects to examine specific methodological questions

The Pilot Study, also known as the Vanguard Study, will run for 21 years, enroll about 4,000 families, and precede Main Study activities by about 3 years so that every aspect of the Main Study is field tested prior to scale up and implementation.
The NCS incorporates an informatics approach that is standards-based

- During the Vanguard phase of the NCS, multiple informatics platforms and tools are being field-tested to determine their performance characteristics.

- This approach entails the use of NCS specifications to which each potential informatics platform must comply from data collection through data submission.

- Use of specifications complements an interoperable approach that allows support for common interfaces and data exchange specifications.

- Specifications are informed by several sources:
  - Data Documentation Initiative (DDI)
  - Clinical Data Acquisition Standards Harmonization (CDASH)
  - CDISC Operational Data Model (ODM)
  - ISO 11179 / 21090
  - CRoss-Industry Standard Process for Data Mining (CRISP-DM)
The NCS uses standard data elements to evaluate study operations

- The NCS incorporates a set of defined data elements referred to as Operational Data Elements (ODEs) that allow for the systematic and objective evaluation of how the study is conducted.

- The ODEs can quantitatively describe study operations and costs while providing a basis for continuous improvement in quality and efficiency.

- The NCS developed a catalog of >500 ODEs for various study operations that address such functions as staffing and expenses.

- The NCS continues to refine these ODEs and would like to contribute to the establishment of standards for others using Operational Data Elements.
NCS metadata can support study operations

- Meta data is the data about the data that provides context, organization and consistency
- The NCS is constructing a meta data framework with concepts, definitions and repository that will be interoperable and harmonized with national and international standards

<table>
<thead>
<tr>
<th>Domain Team</th>
<th>Activity</th>
<th>Current Issue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instrument Development</td>
<td>Question library, sourcing and question</td>
<td>Protocols are often large with many instruments and versions which are not</td>
</tr>
<tr>
<td></td>
<td>standardization</td>
<td>easy navigated and require many manual steps in creating new instruments.</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>Analysis of recruitment kinetics</td>
<td>ODEs do not always support the creation of time series data sets</td>
</tr>
</tbody>
</table>
The NCS Meta Data Repository (MDR) facilitates the use of content and data

- The NCS Meta Data Repository (MDR) uses a standards-based platform that describes, documents and facilitates the use of study content as well as supporting data analysis.
The NCS supports efforts to harmonize pediatric research concepts

- The NCS is engaged in efforts to harmonize terminology across sub-specialties and developmental stages which requires an understanding of and agreement to use specific definitions for terms and concepts (*Neonatal Terminology Formative Project*)

- The development of a harmonized terminology is a fundamental step toward enabling interoperability and data exchange, which then allows aggregation and comparison of data collected at different times or by different groups, resulting in richer analyses

- The terminology system differs from other terminology systems by incorporating into all concepts a dimension of time and position along a developmental scale to relate concepts to one another

- More information on NICHD Efforts to Harmonize Pediatric Terminology can be found at: [http://nichd.nih.gov/health/clinicalresearch/terminology/](http://nichd.nih.gov/health/clinicalresearch/terminology/)
The NCS supports efforts to harmonize pediatric research concepts
Summary

- The NCS utilizes standards from multiple sources to ensure an open-architected, sustainable and interoperable informatics environment.

- The NCS is field testing in the Vanguard or Pilot phase several tools and platforms concurrently in a systematic fashion to determine performance characteristics.

- The integration of multiple standards and models allows exploration of meta data analyses, operational data elements and project management across all study activities.
For more information on the National Children’s Study…

- Please visit the NCS main website: http://www.nationalchildrensstudy.gov/
- Organizations, groups or individuals that are interested in contributing to the effort or learning more are encouraged to contact:

John Lumpkin, MS, MBA, PMP
Steven Hirschfeld, MD PhD
ContactNCS@mail.nih.gov