THE NATIONAL CHILDREN'S STUDY

HEALTH GROWTH ENVIRONMENT

Meeting the Informatics Challenges of the National Children's Study

William R. Hogan, MD, MS

NCS Research Day

August 24th, 2011

Informatics Challenges

- Capturing numerous and diverse types of data…
- …in a comparable way across:
  - Dozens of study centers nationwide
  - Over two decades of time (at least)
- …and for which there are minimal standards
- …and include types of data for planning the NCS
- Data must be of highest research quality!

Accept challenges, so that you may feel the exhilaration of victory.

-George Patton

Victory Defined

Accelerate the progress of science

- Accomplish objectives of National Children's Study
- Advance the science and practice of informatics
- Software applications of general utility for research of all types

Meeting the Challenges: Point of Embarkation

- No adequate software systems available "out of the box"
  - Must therefore assemble a suite
  - Lack of standards ➔ integration is hard
  - Must adapt components to an ambitious protocol
- Minimal functionality for:
  - Data quality
  - Reporting
  - Extraction of data for submission

A Few Guiding Principles for the Journey

- Open-source systems are:
  - Easier to adapt
  - Easier to integrate
  - Easier to "future proof"
- Division of labor and collaborative intelligence will get us there faster
  - Software
  - Instruments
  - Standards, common definitions, etc.
  - Documentation
  - *Innovation*

The Journey So Far

- We and others have:
  - Scoured the landscape for appropriate open-source applications
  - Started integration and adaptation efforts
  - Worked to build communities that make faster progress than in insolation
- The Program Office has:
  - Paid careful attention to standardization
  - Developed methods to analyze quality of data it receives

In Arkansas

- Open-source applications in use
  - caBIG Central Clinical Participant Registry (C3PR) – *particpant registration and consent information*
  - LimeSurvey – *instruments*
  - Open sYStem for Entity Resolution (OYSTER) – *resolving duplicate address records*
- Significant adaptations to C3PR
  - Move participants back and forth among epochs
  - Address/participant entity resolution to resolve duplicate records
  - Restructuring of consents, “checkboxes”, refusals to participate, etc.
- One FTE required to build and update instruments in LimeSurvey

Arkansas Experience Demonstrates That…

- The use of open-source software has indeed enabled more rapid progress
- Modifications of software to meet NCS requirements has benefited other research, and vice versa
- We have benefited from development done by other study centers

Thoughts on the Roadmap from Here

- Given the pressures to demonstrate results quickly, we should:
  - Work more in collaboration
  - Improve infrastructure for collaboration and communication
  - Improve workflow and processes for collaboration
  - Consider formative research projects to achieve these goals
- We are:
  - Study Centers
  - Program Office
  - Subcontractors of both

There is much to learn here from successful open-source communities

Conclusion

- Understanding the informatics solutions needed for the NCS is a research question in its own right
- We are conducting multiple informatics experiments to answer this question
- The results thus far:
  - The informatics opportunities and challenges are as significant (or more) than we anticipated
  - While no single software solution exists to meet the needs of NCS, there are multiple, open-source models proving success
  - Informatics progress thus far is encouraging, and we can improve the efficacy of our efforts by increased collaboration