# Developing Statistical Standards in an Academic Data Coordinating Center (DCC)

Amy E. Clark, MS
Angie S. Webster, MStat
Richard Holubkov, PhD
J. Michael Dean, MD, MBA

Department of Pediatrics,
University of Utah
Salt Lake City, UT

This work was partially supported by the following grants: U01-HD049934 (NICHD), U01-HL094339 (NHLBI), and U03-MC00008 (EMSC, HRSA)

Society for Clinical Trials
May 16, 2011

## Background on the University of Utah DCC

- Started in 2002 in support of a single, multi-institutional research network and two projects
- Currently provide support for 5 multi-institutional research networks and more than 30 active projects
- Over 30 staff including data center PI, 4 program directors, 9 project managers, 7 clinical data managers, and 9 statisticians

2

## Importance of standards

- Optimize quality, consistency, and efficiency of support provided
- Enhance cumulative knowledge and expertise of staff and facilitate transfer of projects
- Positive impact on retention of current funding streams and ability to successfully compete for new funding

3

## Challenges to development of standards in an academic DCC

- Limited resources and infrastructure
- Limited familiarity with requirements for rigorous management and support of studies
- No industry standards and less uniform regulatory requirements compared to private sector
- Diversity of projects

4

## Basic steps in standards development

5

## Get started

- Identify areas to develop standards in
  - Processes that lend themselves to standardization
  - Areas with greatest potential to make a difference in support provided
- Begin outlining content for standards
  - Seek resources and collaborators with expertise in the areas you want to develop
- Evaluate what is necessary, what is ideal, and what is achievable

6

## Write and refine

- Consider balance between level of detail and usability
- Involve staff directly in process
  - Leverages staff experience and insights
  - Promotes buy in and ownership from those who will be responsible for day-to-day implementation
  - Can result in better product and improved adherence

7

## Implement and evaluate

- Provide comprehensive training as standards are finalized and establish expectations
- Evaluate adherence and track over time
- Provide re-training and re-evaluate / revise content periodically

8

## General principles

- Set resources aside for this task
- Develop, refine, and implement piecewise
  - Most critical standards first
  - Create goals and a timeline to ensure you don't lose momentum
- Keep standards manageable
  - Think 1 inch binder!
  - Helps to ensure familiarity and compliance
  - Facilitates maintenance over time

9

## Statistical standards

10

## Statistical support – core functions

- Study design
- Statistical analysis plan
- Analysis datasets
- Statistical analyses; tables, figures and listings
- Report and/or manuscript writing

11

| Areas of Emphasis for Key Statistical Standards |
|---|
| Statistical analysis plans |
| Database and data elements review |
| Analysis dataset documentation |
| Documentation of analysis decisions |
| Creation of raw datasets |
| Creation of derived analysis datasets |
| General programming standards |
| Quality assurance in statistical programming |
| Storing, accessing, and archiving analysis files |

12

| Areas of Emphasis for Key Statistical Standards |
|---|
| **Statistical analysis plans*** |
| **Database and data elements review*** |
| **Analysis dataset documentation*** |
| **Documentation of analysis decisions*** |
| Creation of raw datasets |
| Creation of derived analysis datasets |
| General programming standards |
| Quality assurance in statistical programming |
| Storing, accessing, and archiving analysis files |

13

| Areas of Emphasis for Key Statistical Standards |
|---|
| Statistical analysis plans |
| Database and data elements review |
| Analysis dataset documentation |
| Documentation of analysis decisions |
| **Creation of raw datasets*** |
| **Creation of derived analysis datasets*** |
| **General programming standards*** |
| **Quality assurance in statistical programming*** |
| **Storing, accessing, and archiving analysis files*** |

14

## Example: quality assurance in statistical programming

15

## Importance

- Ensures consistently high quality and accurate results
- Minimizes risks related to errors in programming or analyses
- Provides structure and thoughtfulness to dataset programming and analyses
- Promotes shared learning among statisticians

16

## Assessment of risk

- Related to the possibility of harm associated with coding or analysis errors
  - Patient care including (when applicable) ongoing participation in research
  - Data center reputation
  - Allocation of resources
- Risk level determines quality assurance procedures required

17

## Assessment of risk

| Risk Level | Definition |
|---|---|
| Low | Results will not be used for important decisions |
| Medium | Results may impact study design, funding decisions, or future directions of research but do not expose DCC to external scrutiny or financial risk |
| High | Results directly inform conclusions for key studies or critical aspects of study design and may result in external scrutiny or have financial implications |

18

## Requirements based on risk assessment

- Low risk – basic programming standards
- Medium risk – basic programming standards plus peer review by a second statistician
- High risk – basic programming standards plus independent dual programming of datasets and results by second statistician

19

## Concluding thoughts

20

## Concluding thoughts

- Begin process of standards development as soon as possible and start with critical components
- Set aside resources to enable development and implementation; leverage staff experience and input
- Provide ongoing training and evaluation of standards to ensure adherence and maintain relevance

21

## Questions

Amy E. Clark, MS
amy.clark@hsc.utah.edu

22