

Data Management in Beantown

ADC Data Core Leaders Meeting
October 2, 2004

Suzette Levenson
BU Alzheimer's Disease Center
Director, Data Coordinating Center,
BU School of Public Health



Supported by NIA U01 P30AG13846

IT ISN'T ROCKET SCIENCE....

IT'S JUST CRITICAL!!

Why Data Management?

- Standardized procedures
 - Highest quality of data
 - Reliable conclusions
-
- ```
graph TD; A[Standardized procedures] --> B[Highest quality of data]; B --> C[Reliable conclusions];
```
- The diagram illustrates a causal chain where standardized procedures lead to the highest quality of data, which in turn leads to reliable conclusions. Each step is represented by a bullet point, and the steps are connected by downward-pointing arrows.

# DATA MANGEMENT 101: AN OVERVIEW

- Organizing the TEAM
- Organizing the SYSTEMS
- Designing individual components
  - Tracking systems (subjects & data)
  - Data collection forms
  - Coding manuals
  - Data entry system
  - Quality control procedures
- Technology decisions
- Documentation

# COLLABORATION

- Clinical experts (neurologists, neuropsychologists, e.g.)
- Data collectors (interviewers, phlebotomist, neuropsychology testers)
- Data managers
- Programmers
- Statisticians

# Quality Control Systems

## Public Health Model

- Prevention
- Detection
- Treatment

“A stitch in time saves nine” - Proverb

# PREVENTION : 6 Steps to Nirvana!

1. Identify & assign tasks, staff
2. Identify technology needs (hardware, software, network connectivity)
3. A protocol which clearly defines study procedures
4. Written procedures (Manual of Operations)
5. Well designed data collection forms
6. Training program

# PREVENTION : BU ADC

- Collaborative effort between the Clinical Core and Data Core to create data forms
  - Piloting forms
  - Training
  - Database designed to easily follow flow of data forms
- Clinical team meets weekly with our Data Manager to review each completed form and any problems flagged for review and remediation
- Secure, password protected SQL database



# QUALITY CONTROL SYSTEMS

## Public Health Model

---

- Prevention
- Detection
- Treatment

# QUALITY CONTROL SYSTEMS

- Low Tech
  - Manually review forms for completeness
    - Legibility
    - Correct skip patterns
    - Correct logic (e.g., dates )
  - “Chanting”
- High Tech
  - Double entry with report comparing 2 entries
  - Automated logic and range checks
    - Reports generated from database
    - Automated “paper trail” - audits

## Hope Study

|    | Autopsy #              | PIII  | FHSID | Form Completed on | Checked?                            | Entered By | Entered On |
|----|------------------------|-------|-------|-------------------|-------------------------------------|------------|------------|
| 1  | <a href="#">A03-37</a> | 10432 |       | 11/24/2003        | <input checked="" type="checkbox"/> | framkatie  | 05/05/2004 |
| 2  | <a href="#">A03-40</a> | 10467 |       | 12/30/2003        | <input checked="" type="checkbox"/> | framkatie  | 05/05/2004 |
| 3  | <a href="#">A03-41</a> | 10535 |       | 12/31/2003        | <input checked="" type="checkbox"/> | framkatie  | 05/05/2004 |
| 4  | <a href="#">A03-45</a> | 10516 |       | 01/20/2004        | <input checked="" type="checkbox"/> | framkatie  | 05/05/2004 |
| 5  | <a href="#">A03-55</a> | 10500 |       | 03/31/2004        | <input checked="" type="checkbox"/> | framkatie  | 05/05/2004 |
| 6  | <a href="#">A03-58</a> | 10444 |       | 04/01/2004        | <input checked="" type="checkbox"/> | framkatie  | 05/05/2004 |
| 7  | <a href="#">A04-02</a> | 10439 |       | 04/26/2004        | <input checked="" type="checkbox"/> | DDixon     | 06/17/2004 |
| 8  | <a href="#">A04-05</a> | 10486 |       | 07/12/2004        | <input checked="" type="checkbox"/> | DDixon     | 08/08/2004 |
| 9  | <a href="#">A04-07</a> | 10502 |       | 07/26/2004        | <input checked="" type="checkbox"/> | DDixon     | 08/08/2004 |
| 10 | <a href="#">A04-08</a> | 10560 |       | 07/27/2004        | <input checked="" type="checkbox"/> | DDixon     | 08/08/2004 |
| 11 | <a href="#">A04-12</a> | 10554 |       | 08/02/2004        | <input checked="" type="checkbox"/> | DDixon     | 08/08/2004 |
| 12 | <a href="#">A04-13</a> | 10537 |       | 08/03/2004        | <input checked="" type="checkbox"/> | DDixon     | 08/08/2004 |
| 13 | <a href="#">A04-15</a> | 10530 |       | 08/18/2004        | <input type="checkbox"/>            | DDixon     | 09/09/2004 |
| 14 | <a href="#">A04-16</a> | 10570 |       | 08/19/2004        | <input type="checkbox"/>            | DDixon     | 09/09/2004 |
| 15 | <a href="#">A04-17</a> | 10389 |       | 08/19/2004        | <input type="checkbox"/>            | DDixon     | 09/09/2004 |
| 16 | <a href="#">A04-18</a> | 10576 |       | 08/20/2004        | <input type="checkbox"/>            | DDixon     | 09/13/2004 |
| 17 | <a href="#">A04-21</a> | 10524 |       | 08/24/2004        | <input type="checkbox"/>            | DDixon     | 09/09/2004 |

# Example: Automated Double Entry Report

## Discrepancy Between 1st & 2nd Data Entry

FORM: Neuropathology: Section A

Data Updated on: 12/17/2003

NPID: 136

|          | Field Name: | 1st entered By | First Entry | Updated Value* | Updated By |
|----------|-------------|----------------|-------------|----------------|------------|
| <b>1</b> | DOA         | j136           | 7/27/2003   | 7/26/2003      | jsy        |
| <b>2</b> | PM          | j136           | 04:30       | 41:30          | jsy        |
| <b>3</b> | SEX         | j136           | 2           | 1              | jsy        |

Note: This report only shows the fields on which the second entry was different from 1st entry.  
Updated Value\* is the final value stored in the database

Query Type: Question

| Location                            | Date<br>Time     | Query Details                                                                                                                                                             | Response |
|-------------------------------------|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|
| Form#: 8                            |                  |                                                                                                                                                                           |          |
| Question: F8ADATE<br>Query #: 26084 | 2/28/99<br>03:32 | This transfusion occurred before the first dose date. Please either delete this transfusion or correct the transfusion date, and inform us of the corrective action used. |          |

# SITE VISITS

- Walk-through procedures
- Maintenance of local data systems
- Audit sample of records (5-10%)
  - Compare data collection form data with source documents and computer database output
- Examine consent forms
- A complete audit of first few subjects

# STATISTICAL APPROACHES

- Inspection of data by
  - Center
  - Diagnosis
  - Repeated measures within subjects
  - Secular trends
- Reliability - agreement between 2 observers
- Validity - internal consistency (range & logic checks)

# STATISTICAL TOOLS

- Stem & Leaf plots
- Frequency plots
- Variance
  - too skewed
  - too little



# People Resources



# How To Accomplish QC & QA?

- Time
- Staff With Expertise
- Technology
- \$\$MONEY