Data Management and Database Tools In Research

Presenter and Translator from English into Russian: Elena O’Donald, PhD, MCP, CCNA
Ph.D. in Biological Sciences, MS in Computer Engineering

Bi-National Health Risk Seminar Series
Albuquerque, New Mexico
November 19, 2014
Acknowledgements

The NBCS Project Research Team:

Many thanks to the Navajo Birth Cohort Study (NBCS) Research Team: Johnnye Lewis, Becky Smith, Miranda Cajero, Esther Erdei, Jennifer Ong, Carla Chavez, Bernadette Pacheco, Chris Shuey, Jackie Morgan, Joe Hoover, Curtis Miller, Debra MacKenzie, Carol (CJ) Laselute, and Lauren Hund

Funding Acknowledgement:

CDC’s (Centers for Disease Control) Award Number for the NBCS: CDC U01 TS000135-05
Presentation Overview

Data Storage and Management Tools:
- Microsoft Access Database Tool
- RedCAP Online Survey Tool

MS Access Databases for the NBCS project:
- NBCS BioMonitoring Database
  - Parsing Data
  - Generating Letters to Participants
- NBCS Home Environmental Assessment Database
  - Data Collection Form

Challenges in Data Management
Data Storage and Management Tool

RedCAP Online Survey Tool

We use it primarily for: Data Entry and Data Storage

Advantages:

Security Policy Compliant
Accessible Online

No special database software needed to be installed

Fast Development of basic data entry forms and surveys

Does not require database development and programming skills to make surveys

Disadvantages:

This is not a Relational Database Development Tool

Data Linking Capabilities are extremely limited
This results in irrational data structure design

Form Development is Highly Restricted: limited availability of basic object types
Limited Form Design does not provide good visibility of data on larger forms

Highly Limited Query Capabilities
Design of Query Building Tools is not user friendly

Difficult in Managing Data
Extremely difficult in doing Data Quality Control
Data Entry: Event Grid

The grid below displays the form-by-form progress of data entered into the project for one particular events. You may click on the colored buttons to access that form for that event. If you wish, you may navigating to the Define My Events page.

Study ID SKL034

<table>
<thead>
<tr>
<th>Data Collection Instrument</th>
<th>Enrollment (1)</th>
<th>Dose 1 (2)</th>
<th>Visit 1 (3)</th>
<th>Dose 2 (4)</th>
<th>Visit 2 (5)</th>
<th>Dose 3 (6)</th>
<th>Visit 3 (7)</th>
<th>Final visit (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td>Green</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Contact Info</td>
<td>Green</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Data</td>
<td>Green</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visit Lab Data</td>
<td>Green</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Patient Morale Questionnaire</td>
<td>Green</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visit Blood Workup</td>
<td>Green</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Visit Observed Behavior</td>
<td>Green</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completion Data</td>
<td>Green</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Completion Project Questionnaire</td>
<td>Green</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Relational Database Development Tool

Microsoft Access
It can be used for: Entering, Storing, Querying, Managing, and Reporting Data

Advantages:
Security Policy Compliant
Database can be made Accessible Online (with limitations)
Medium Speed of Database Development
Flexibility in Designing: Database Structure, Tables, Forms, Queries, and Reports
Easy to Link Data and develop Complex Queries and Reports
Flexible Form Design allows good visibility of data on larger forms
Good for Data Entry, Management, and Quality Control
User Friendly Environment for Database and Visual Basic Code Development

Disadvantages:
Installation of MS Access is required
For developer: MS Access database development skills are required
For developer: MS Visual Basic Programming skills are desired
Limited Form Design Capabilities for the Online version of Database
MS Access Databases for the NBCS project

**NBCS BioMonitoring Database**

We currently use it for:

**Parsing Data**
- CDC – Blood and Urine Test Results
- RedCAP – Participant Information

**Processing Data**
We calculate 50-th and 95-th percentiles for the NBCS sample

**Generating Letters to Participants**

* * * Let’s now look at some of these forms * * *
MS Access Databases for the NBCS project

NBCS
Home Environmental Assessment Database

It is currently in development

It will be used for:
Collecting Data on the Field
Data Quality Control
Generating Letters to Participants

* * * Let’s now look at the Data Entry Form * * *
Some of the Challenges that we faced while Managing Data
Some organizations:
Keep changing format of data given to us,
They inconsistently return processed data

* * *

Majority of Incoming Data is *not* normalized

* * *

There is no standardized process of Data Exchange between different organizations
Challenge of using computers

Majority of data collection people who lived and worked at remote locations had no experience of using computers.

The off-site employees had to learn basic computer skills before starting entering the project’s data.

We provided training sessions to help them and they practiced on their own until they started feeling comfortable using computers.
Incompatible Confidentiality Protection that is utilized by different organizations results in the following...
HIPAA and HSC have their own firewalls

That’s why our users sometimes have to work on the two separate computers at the same time, if they want to link data located at both sites.

Notes: HIPAA = U.S. Department of Health and Human Services; HSC = UNM Health Sciences Center
Funding is sourced to various organizations that have different policies, needs, and goals. That’s why our users sometimes struggle to have the work done by partner organization’s employees on time and with required degrees of accuracy and consistency.