Lifecycle Processes to Insure Quality of Blaise Interview Data

Linda Gowen, Westat
Pat Clark, Westat
This presentation reviews some of the Westat tools used to facilitate processing at the different stages throughout the lifecycle of a Blaise project, with a focus on the impact on data quality.
Survey Lifecycle Overview

• The major lifecycle steps are:
  • Design
  • Development
  • Data Collection
  • Data Preparation
  • Data Delivery

• Westat tools are used at every stage to improve data quality.
Lifecycle Processes

**Lifecycle Steps**
- Design
- Development
- Data Collection
- Backend/Data Preparation
- Data Delivery

**Stages**
- Design Survey Instruments
- Develop CAI Questionnaire Specifications
- Develop Blaise System
- Test Blaise System
- Pretest Blaise System
- Deploy Blaise System
- Edit and QA Blaise Data
- Create Codebooks
- Deliver Data and Documentation/Prepare For Follow-up

**Westat Systems**
- SpecWriter
- Blaise
- BFOS
- Blaise Editing System
- Westat Meta Data Delivery System

**Date:** 04/09/2008
Lifecycle Processes

**Lifecycle Steps**
- Design
  - Design Survey Instruments
- Development
  - Develop CAI Questionnaire Specifications
  - Develop Blaise System
  - Test Blaise System
  - Pretest Blaise System
- Data Collection
  - Deploy Blaise System
- Backend/Data Preparation
  - Edit and QA Blaise Data
  - Create Codebooks
- Data Delivery
  - Deliver Data and Documentation/Prepare For Follow-up

**Stages**
- Design Survey Instruments
- Develop CAI Questionnaire Specifications
- Develop Blaise System
- Test Blaise System
- Pretest Blaise System
- Deploy Blaise System
- Edit and QA Blaise Data
- Create Codebooks
- Deliver Data and Documentation/Prepare For Follow-up

**Westat Systems**
- SpecWriter
- Blaise
- BFOS
- Blaise Editing System
- Westat Meta Data Delivery System

04/09/2008
Stage 1 - Requirements Analysis

- The goal of this stage is to define both the Blaise and non-Blaise requirements for the system.

- It is helpful to prepare a project data flow diagram showing:
  - All sources providing data into the Blaise Systems
  - All data required to feed external systems
  - All data transformations
Stage 2 - CAI Specifications

The CAI specifications document includes the “complete” hard copy questionnaire typically needed for regulatory compliance, with additional detailed instructions, necessary for programming the CAI system.

<table>
<thead>
<tr>
<th>Question#</th>
<th>Field Name</th>
<th>Question Text</th>
</tr>
</thead>
</table>
| Q23.1     | Cough_AM   | Now I am going to ask you about respiratory symptoms that you may have experienced in the past year. Do you usually cough on waking up, or first thing in the morning?
|           |            | 0. NO  
|           |            | 1. YES  
|           |            | Don't Know.  
|           |            | Refused.   |
| Q23.2     | Cough_PM   | Do you usually cough during the rest of the day or at night?
|           |            | 0. NO  
|           |            | 1. YES  
|           |            | Don't Know.  
|           |            | Refused.   |
| Q23.3     | PastYrCough_Plus3Mos | During the past 12 months, have you had this cough on most days for three months or more?
|           |            | 0. NO  
|           |            | 1. YES  
|           |            | Don't Know.  
|           |            | Refused.   |

IF Question Q23.1(Cough_AM) <> Yes
AND Question Q23.2(Cough_PM) <> Yes
GOTO Question Q23.3(PastYrCough_Plus3Mos)
Stage 3 - Blaise Development

- Programmers reference documents
  - Project data flow diagram
  - CAI specifications
Blaise Development Key Considerations

- Proper use of data to be used as foreign keys in the creation of blocks/tables
- Review of complicated programming
- Testing and QC at all stages of instrument development including verification of the data integrity after both forward and backward navigation
- Declaring fields to be compatible with external systems
- Review relationship of question numbers to Blaise variable names
- Save needed intermediate variables
Blaise Development Operational Considerations

- Identify data for external systems performing quality assurance and monitoring.
- Advance planning can assure the database structure provides for the easy extraction of data.
Stage 4 - Internally Testing the Blaise System

- A specialized testing team ensures a level of quality
- CAI Specifications are tested
- Issue Tracking
- Feature by feature validation
- Functional testing
- Testing is done using all systems
- All testing is based on workflow
- Regression testing
Stage 5 - Pretest for Field Readiness

- Testing is also performed by the project staff, data management staff, and field staff.

- Role plays should simulate data collection scenarios.

- The goal is to have each real condition tested for all the responses available in the instrument.

- Files with the delivery layout are built from the mock data and frequencies run to verify that the system is capturing the data to meet the research and operational objectives.
[HAVE TRAINEE READ]

WaterSource.

Let’s talk about your current home in Shaftsbury. You said that you lived here from 1995 to the present. During this period, what source or sources supplied water to this home?

[INTERVIEWER: SHOW CARD B3-1. SELECT ALL THAT APPLY.]

[INTERVIEWER: BROWSE WATER SOURCE (PUBLIC OR PRIVATE) OR CATEGORY.]

A PRIVATE WELL OR SPRING ................ (10)
A COMMUNITY WELL OR SPRING
SERVING AT LEAST 25 PEOPLE OR IS
IMPLANTED ................... (11)
A PUBLIC OR MUNICIPAL WATER
SUPPLY .................. (12)
ANOTHER SOURCE OF WATER
(SPECIFY) ................ (13)

[RESPONDENT]: “I've got city water and I also have a private well.”

NumberWellsSprings.

How many different private wells or springs supplied water to this home?

[INTERVIEWER: ENTER “0” FOR FIVE OR MORE PRIVATE WELLS OR SPRINGS.]

1 1

[RESPONDENT]: “One”
Lifecycle Processes

**Lifecycle Steps**
- Design
- Development
- Data Collection
- Backend/Data Preparation
- Data Delivery

**Stages**
- Design Survey Instruments
- Develop CAI Questionnaire Specifications
- Develop Blaise System
- Test Blaise System
- Pretest Blaise System
- Deploy Blaise System
- Edit and QA Blaise Data
- Create Codebooks
- Deliver Data and Documentation/Prepare For Follow-up

**Westat Systems**
- SpecWriter
- Blaise
- BFOS
- Westat Meta Data Delivery System

Date: 04/09/2008
### Cases:

<table>
<thead>
<tr>
<th>Case ID</th>
<th>Respondent Name</th>
<th>Disposition</th>
<th>Disposition Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>901304B</td>
<td>George Washington</td>
<td>EL-Eligible</td>
<td>Thu. 5/25/2006 10:16 AM</td>
</tr>
<tr>
<td>901309B</td>
<td>John Doe</td>
<td>EL-Eligible</td>
<td>Thu. 5/25/2006 12:38 PM</td>
</tr>
<tr>
<td>901310B</td>
<td>Mary Clark</td>
<td>EL-Eligible</td>
<td>Wed. 5/31/2006 2:32 PM</td>
</tr>
<tr>
<td>901403A</td>
<td>Dennis Miller</td>
<td>EL-Eligible</td>
<td>Wed. 5/31/2006 3:13 PM</td>
</tr>
<tr>
<td>901509A</td>
<td>David Williams</td>
<td>EL-Eligible</td>
<td>Fri. 6/2/2006 11:07 AM</td>
</tr>
</tbody>
</table>

### Tasks:

<table>
<thead>
<tr>
<th>Screener</th>
<th>Baseline</th>
<th>Quarterly 1</th>
<th>Quarterly 2</th>
<th>Quarterly 3</th>
<th>Quarterly 4</th>
<th>Quarterly 5</th>
<th>Quarterly 6</th>
<th>Quarterly 7</th>
<th>Followup</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO-Complete</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
BFOS provides projects with:

- **Interviewer Management System** - A laptop-based application that allows interviewers to review and change assigned case status, launches the Blaise CAPI instrument, and synchronizes the management and CAPI data with the home office.

- **Supervisor Management System** - This is the web-based application used by field supervisors to view and manage the work of interviewers, including case assignment and transfer between interviewers, review and adjustment of case status, and other administrative and reporting functions.

- **Home Office Management System** - This provides functions needed by home office staff to load and manage the study sample information and to maintain information about the field staff and their assignments.
Planning Prevents Unexpected Results
Lifecycle Processes

**Lifecycle Stages**

- Design
  - Design Survey Instruments
  - Develop CAI Questionnaire Specifications
  - Develop Blaise System
  - Test Blaise System
  - Pretest Blaise System
  - Deploy Blaise System

**Backstage/Data Preparation**

- Edit and QA Blaise Data
- Create Codebooks

**Data Collection**

- Deploy Blaise System

**Data Delivery**

- Deliver Data and Documentation/Prepare For Follow-up

**Westat Systems**

- SpecWriter
- Blaise
- BFOS
- Westat Meta Data Delivery System

04/09/2008 20
Stage 7 - Editing Blaise Data

An array of systems are available to support Blaise editing.

- Editing system
- Extract system
- Comment system
- Data decision log
- Data and interview validation
- QA reconciliation systems
- Study management systems
- Metadata delivery system
Basic steps of the editing process are:

- Combine data into a master database
- Maintain Blaise interview in interviewing and data editing modes
- Display comments
- Editors may make changes to data.
- Decisions documented.
- Other, specify fields coded.
- Frequencies reviewed
- Reconciliation completed.
### Question:
During the past 12 months, have you had this cough on most days for three months or more?

#### Options:
- 0. NO
- 1. YES

**Remark:**
more prevalent in spring and fall

<table>
<thead>
<tr>
<th>Question</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q20.1 Cough</td>
<td>1</td>
</tr>
<tr>
<td>Q20.2 Cough</td>
<td>0</td>
</tr>
<tr>
<td>Q20.3 FastYr</td>
<td>6</td>
</tr>
<tr>
<td>Q20.4 Cough</td>
<td>15</td>
</tr>
<tr>
<td>Q20.5 Cough</td>
<td>0</td>
</tr>
<tr>
<td>Q20.6 Cough</td>
<td>1</td>
</tr>
<tr>
<td>Q20.7 FastYr</td>
<td>0</td>
</tr>
</tbody>
</table>
Frequency Review

- As a final check of the data files themselves, the data are extracted from the final edited tables and frequency reviews are performed.
Quality Assurance

- Redundant data.

- QA system
  - Report discrepancies
  - Secure
  - Reports on Editing system data, SMS data, other project source data

- Documentation of resolutions recorded in Data Decision Log.
Stage 8 - Create Codebooks

- The codebook is an important part of documentation.
- Our codebook generation system was designed around the metadata output from the Blaise instrument.
- Data elements are versioned.
- All derived data elements have the associated derivation details present in notes.
### Field Name: CAPI_AGE (7672)
#### Data Element Name and (DE Number):
- CAPI

#### Element Source:
- Number: 3

#### Source Details:
- Form ID
- Form Section
- Item ID
- Date Type/Length
- Position or Column Start: End
- Format

#### Derivation Details:
- CAPI_Age = Consents.Study.CAPIDemographics.DOB.DOB_month, DOB_Day, DOB_year

<table>
<thead>
<tr>
<th>Value</th>
<th>Definition (Skip)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0-29</td>
<td>= Less than 30</td>
<td></td>
</tr>
<tr>
<td>30-100</td>
<td>= Over 30</td>
<td></td>
</tr>
</tbody>
</table>

Total = 387

---

### BOX 1

IF PARTICIPANT IS LESS THAN 30 YEARS AT TIME OF INTERVIEW, GO TO WF_CurrentWeight.(CAPI_AGE < 30) Go to WF3/WF_CurrentWeight

### Field Name: WF_HighAge2Feet (2760)
#### Data Element Name and (DE Number):
- CAPI
- WF_v7
- Weight
- WF_v1

#### Element Source:
- Number: 3

#### Source Details:
- Form ID
- Form Section
- Item ID
- Date Type/Length
- Position or Column Start: End
- Format

#### Item Text: How tall were you at age 21? [INTERVIEWER: ENTER FEET]

<table>
<thead>
<tr>
<th>Value</th>
<th>Definition (Skip)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-6</td>
<td>= FEET</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>= DON'T KNOW</td>
<td></td>
</tr>
</tbody>
</table>

Total = 383

---

### Field Name: WF_HighAge2Inch (2741)
#### Data Element Name and (DE Number):
- CAPI
- WF_v7
- Weight
- WF_v1

#### Element Source:
- Number: 3

#### Source Details:
- Form ID
- Form Section
- Item ID
- Date Type/Length
- Position or Column Start: End
- Format

#### Item Text: HOW TALL WERE YOU AT AGE 21? [INTERVIEWER: ENTER INCHES]

<table>
<thead>
<tr>
<th>Value</th>
<th>Definition (Skip)</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>= NONE</td>
<td>45</td>
</tr>
<tr>
<td>1-11</td>
<td>= INCHES</td>
<td>308</td>
</tr>
</tbody>
</table>

Total = 383

---

Prepared By: Wessell, Rockville, MD

The Kidney Cancer Study - Codebook

Page 2
Lifecycle Processes

**Lifecycle Steps**
- Design

**Stages**
- Design Survey Instruments
- Develop CAI Questionnaire Specifications
- Develop Blaise System
- Test Blaise System
- Pretest Blaise System
- Deploy Blaise System
- Edit and QA Blaise Data
- Create Codebooks

**Westat Systems**
- SpecWriter
- Blaise
- BFOS
- Blaise Editing System
- Westat Meta Data Delivery System

**Data Delivery**
- Deliver Data and Documentation/Prepare For Follow-up

04/09/2008 29
Stage 9 - Data Delivery

Westat has developed a Data Delivery Metadata System which supports delivery associated with datasets, such as:

- Raw data
- Research datasets
- Analytic datasets
- Restricted use datasets, and
- Public Use files.
Summary

- Managing aspects of the processing extending across multiple lifecycle stages is critical to data quality.
- Feedback from all aspects of the survey need to be considered, and iterative processing loops are necessary.
- Lifecycle tools can be implemented in multiple project environments and customized to the needs of individual surveys.
- Active participation from all team members assure that differing perspectives are represented.
- Quality control becomes continuous from planning through delivery, ensuring accurate and efficient data collection and processing.