

# Development of a Data-Driven Data Collection System for Health Examination Surveys

Debra S. Reed-Gillette, MS and Lewis Berman, PhD

## Introduction

### What is NHANES?

- Since 1959 the National Health and Nutrition Examination Surveys (NHANES) have collected data which are best obtained by direct physical examination, clinical and laboratory tests, personal interviews, and related measurement procedures.
- These data are used to establish baseline estimates and to monitor trends in disease and health conditions.



## NHANES Data Collection Tools 1959 - 2002

### Past NHANES data collection tools

- 1959-1984
  - PAPI (Paper and Pencil Interviews)
  - Paper data collection forms for biomedical data
- 1984-1998
  - PAPI
  - CAPI (Computer Assisted Personal Interviews) developed in Cheshire™
  - Paper data collection forms for biomedical data
  - Oracle™ forms
- 1999-2002
  - CAPI and ACASI (Audio-Computer Assisted Self Interviews) developed in Blaise™
  - PowerBuilder™

## What was needed

### In 2002, DHANES needed to quickly develop a framework for Community Based Studies and specifically for NYCHANES

- Requirements for Community Based Studies
  - Short survey time-frame
  - Limited staff, funds, resources
  - Limited IT support
  - Ease backend challenges for inexperienced staff
  - Shrink-wrap NHANES / CHANES architecture
- Limitations of NHANES in 2002
  - Systems in use would require major rework with modifications
  - Inflexible
  - No repurposing of existing Metadata or Systems
  - Slow for development
  - Costly to change and maintain

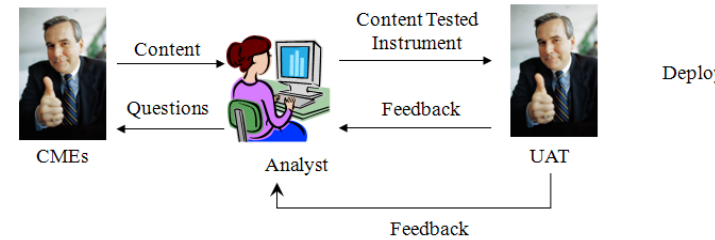
## Options Examined

### No single solution available in 2002. Two options were examined.

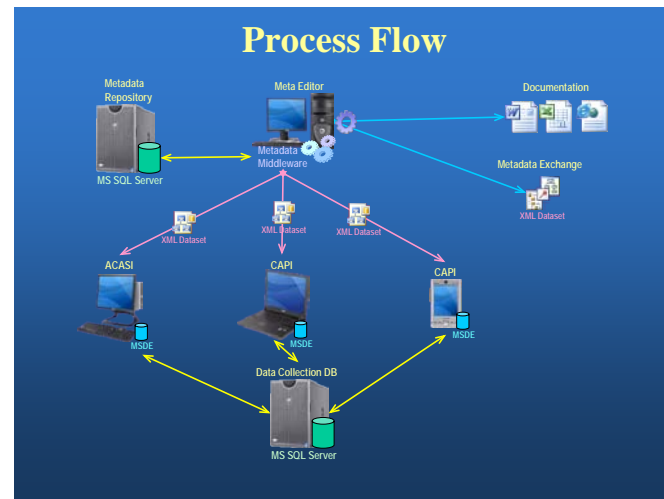
- Blaise™ with PowerBuilder™
  - Large base of users in government
  - Familiarity
  - Some software available for reuse
  - Needed major modifications
  - Not easy to learn
  - Costly for licensing for small communities
- Build our own in-house
  - Customizable
  - In-house development costs lower than licensing fees and modifications to existing systems
  - Reusable metadata for multiple purposes/communities
- No other COTS/GOTS available in 2002 that would be a single solution

## Development

- Environment: MS Visual Studio .NET
- Programming Language: C#
- Database: MS SQL Server, MSDE, XML
- Time: < 1.5 Person-Years
- Team: Sr. Architect, 2 developers, 2 Testers, 1 Analyst
- Initial Testing: Unit, Integration, Acceptance, Full-scale Pilot Tests



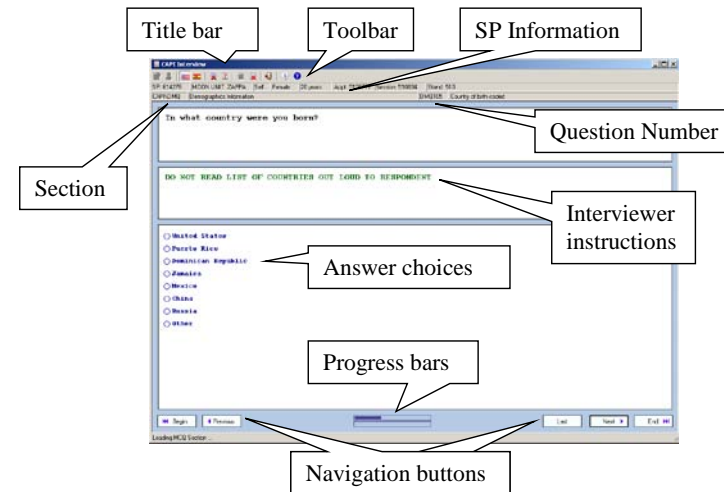
## Process Flow of Developed System



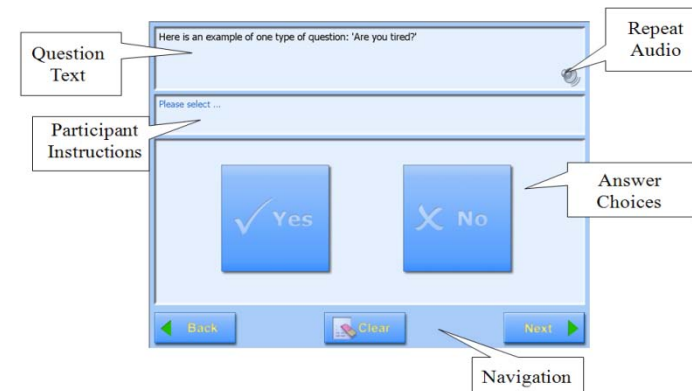
## Features of Initial Developed System

- Fits into NHANES architecture
- Hierarchical metadata model
- XML Database / Relational Database
- Skip pattern / Check item integration, look-up tables, ...
- Reconciliation
- Network and stand-alone mode
- Completion status at section and questionnaire level
- Tutorials
- Multi-language support
- 50 attributes describing individual items
- Media file support
- Version control
- Substitution pattern support

## Computer Assisted Interviewing Interface

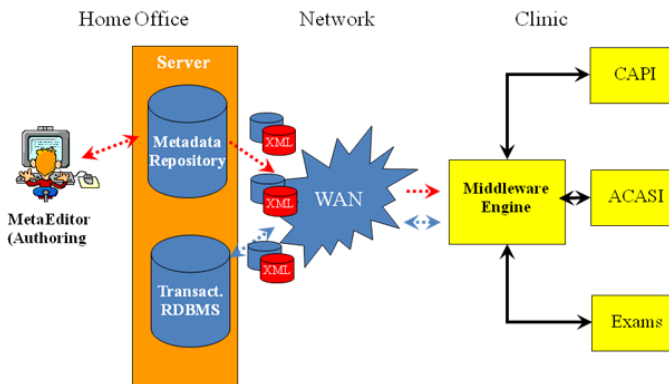
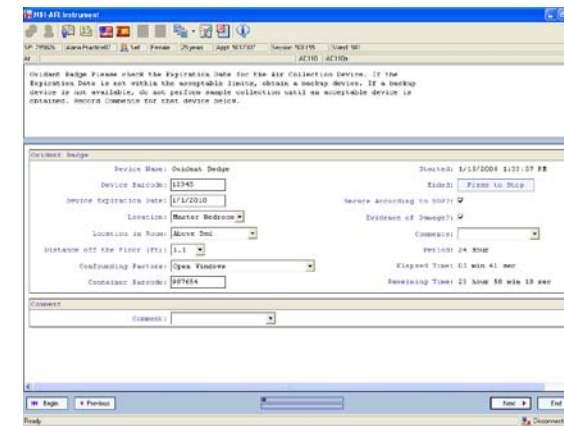


## Audio Computer Assisted Interviewing Interface



## Advanced Feature Development 2005-2009

- Biomedical Equipment Integration
- Advanced Documentation
- Questionnaire Mapping and Logic Flow
- Text to Speech
- Additional Data Collection Field Controls
- Automated Field Deployment



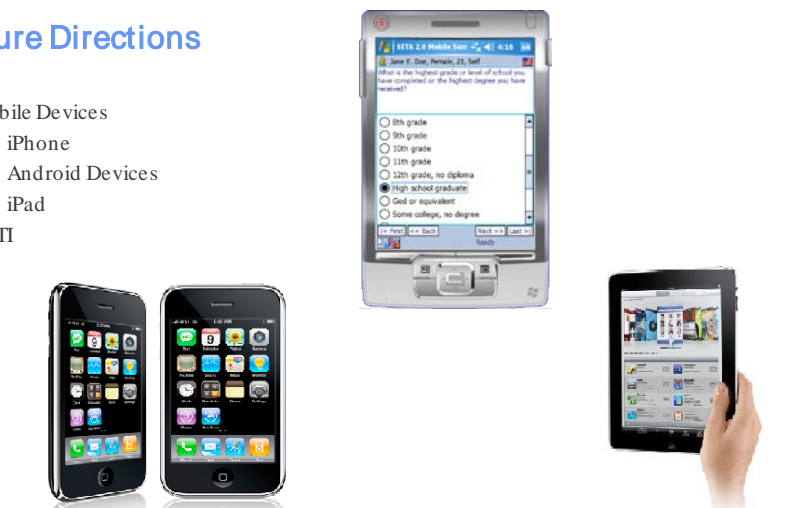
## Multiple Application Deployments

- NYCHANES 2003-2004
- NHANES 2005-present
  - CAPI
  - ACASI
  - Exhaled Nitrous Oxide
  - Spirometry
  - Pubertal Maturation
- Children's National Medical Center 2007
  - Pubertal Maturation Validation Study
- National Children's Study 2006-present
  - Enhanced and expanded to become DCAS
  - The Data Collection Application Suite (DCAS) is a government-owned suite of applications for health studies support. <http://www.dcasproject.org/>



## Future Directions

- Mobile Devices
  - iPhone
  - Android Devices
  - iPad
- CATI



Debra S. Reed-Gillette, MS

NHANES/Operations Branch  
Computer Scientist  
voice: 301.458.4456  
email: DReed-Gillette@cdc.gov

Lewis Berman, PhD

NHANES  
Deputy Director  
voice: 301.458.4096  
email: LBerman@cdc.gov

<http://www.cdc.gov/nchs/nhanes.htm>

