Multiple Overlapping Waves - Challenges in Supporting Blaise Instruments Simultaneously for Four Waves of Data Collections

Lilia Filippenko, Mai Wickelgren, Venkat Yetukuri, Sridevi Sattaluri, and Preethi Jayaram, RTI International

1. Introduction

At RTI International we conduct a number of longitudinal studies where data for subsequent waves is collected after completing a previous one. In contrast, the Work, Family and Health Study (WFHS) must conduct several waves simultaneously. It is a longitudinal study collecting data from individuals in the workplace and in the home at baseline and at 6, 12, and 18 months post-baseline using both in-person and telephone computer-assisted interviews, basic health measures (height, weight, and blood pressure), blood collection, saliva collection, and Actigraphy (measuring sleep quality). Spouses or partners and selected children (between the ages of 9 and 17) of participating employees were recruited into the study during the baseline period. Interviews in the workplace are conducted on a rolling schedule; at one point 20 Blaise interviews were in production in the field for all four waves.

To deal with so many instruments, we developed for WFHS an application in .Net that uses the Blaise Component Pack (BCP) to maintain SQL Server tables for a subset of fields from the Blaise databases. The application utilizes SQL Server stored procedures to create up-to-date preload information for follow-up Blaise interviews, set statuses in a Control System for monitoring purposes, and populate reportable statuses of blood collection, saliva collection, and Actigraphy. The clients are able to view reports, download encrypted files, and to record blood spot counts, saliva receipts and tracking statuses through a secure web portal.

In addition to standard reports, about one hundred custom reports were developed to monitor data collection. A Field Management System and Control System were modified to allow reports to be viewed by wave.

The paper will describe how new and existing applications developed at RTI International help make data collection for WFHS efficient and accurate.

2 Overview of Work, Family and Health Study

The WFHS is composed of interdisciplinary research teams from the University of Minnesota, Penn State University, Harvard University, Portland State University, Michigan State University, Kaiser Permanente's Center for Health Research, and RTI International. The WFHS is studying the impact of workplace programs and policies and how they affect employees' work, health and family members' health and well-being.

2.1 Computer-Assisted Interviews

WFHS data collection includes survey and biomarker data collection led by RTI International, and Harvard University, a telephone-based Daily Diary data collection of participating families led by Penn State University, and qualitative interviews led by the University of Minnesota, Portland State University, Michigan State University.

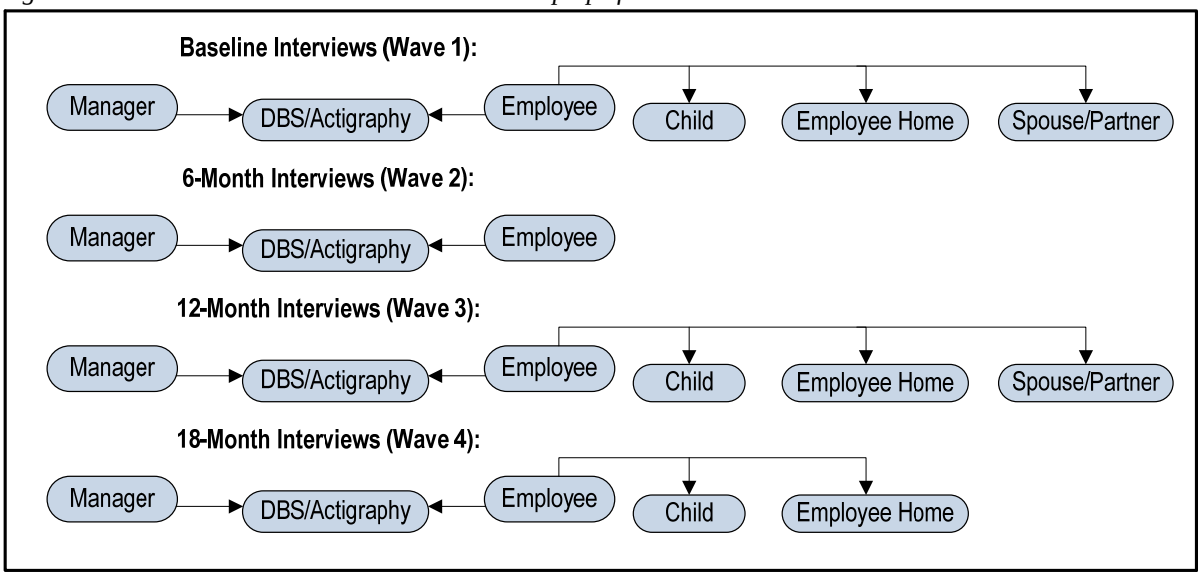
Data is collected from employees and managers within two industries and 56 sites about work experiences (including opinions about work, amount and type of work, job demands), health

(including physical health symptoms, medical care, emotional well-being, and health behaviors), and family structure and relationships.

For four waves of data collection RTI International developed 20 Blaise instruments to be conducted by Field Interviewers on laptops:

- Computer Assisted Personal Interviewing (CAPI) worksite interviews with employees and managers
- DBS (dried blood spots)/Actigraphy (collecting data via an actigraphy watch) for employees and managers
- CAPI home interviews with employee and eligible child with Daily Diary study recruitment and enrollment
- Spouse/Partner telephone interview (administered by the field interviewer)

Figure 1: Blaise Interviews on Field Interviewer Laptops for 4 Simultaneous Waves



In addition to the described above Blaise instruments, two Blaise instruments were developed to conduct interviews with a subset of employee and managers:

- Telephone Verification interview to check the quality of FI work during any in-person interviews.
- Telephone Attriter interview with employees and managers who leave the workplace as of the 6, 12, or 18 month follow-up visit.

These telephone interviews are conducted at the RTI Call Center Services (RTI-CCS) facility.

2.2 WFHS Systems

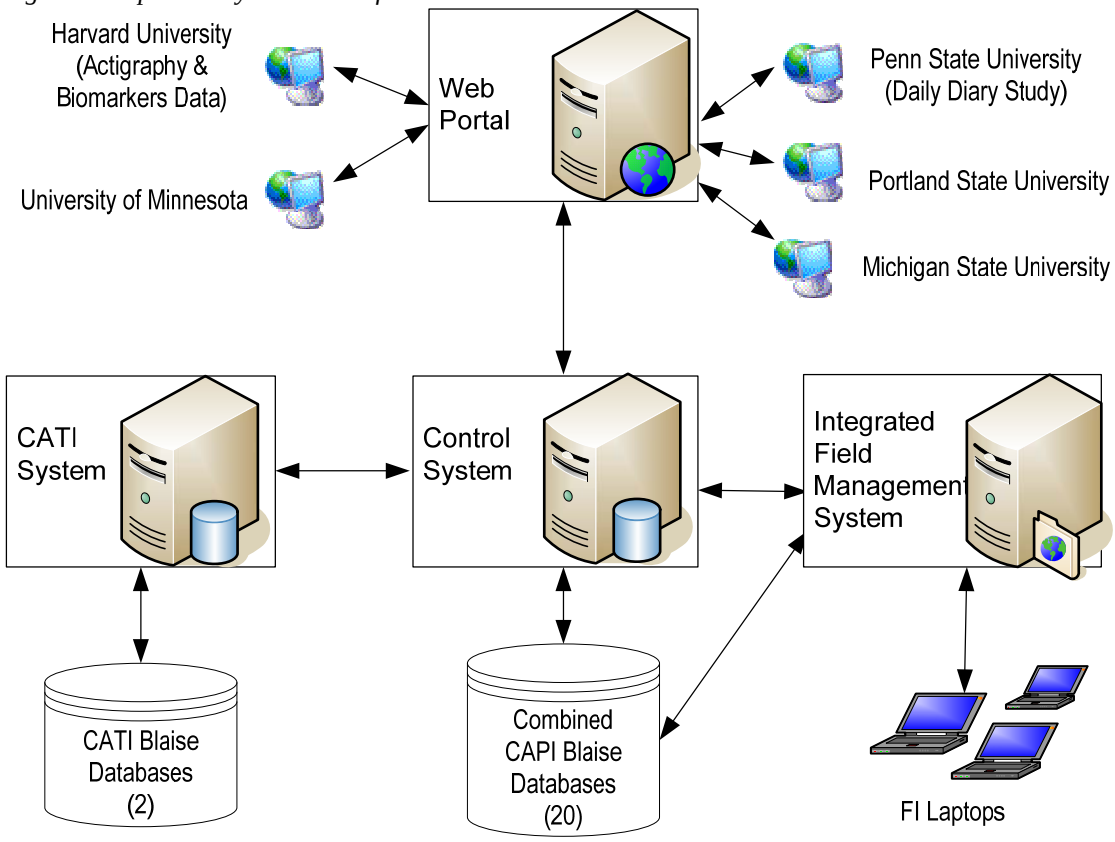
During the development stage of the study, the following standard RTI systems were adopted to satisfy the requirements of WFHS:

- Case Management System (CMS) - used on Field Interviewer (FI) laptops to allow them to update case status, enter comments, launch Blaise instruments, and synchronize the status of cases with a centralized SQL server database.
- Integrated Field Management System (IFMS) - a web application used to assign and transfer cases between FI laptops and RTI.
- General Survey Control System (GSCS) - a web application on RTI's internal secure network used by authorized staff to monitor the flow of data from the start of data collection through the creation of data files for analysis and delivery.

- Computer Assisted Telephone Interviewing (CATI) CMS - used on RTI's secure network to conduct telephone interviews.

To allow all researchers participating in WFHS to monitor data collection and receive up-to-date information from the CAPI interviews, a secure web portal was set up by RTI. During nightly processing at RTI, encrypted files are posted for the Daily Diary study and the Actigraphy study along with status reports about CAPI interviews in the field.

Figure 2: Top Level Systems Used for WFHS



Due to the complexity of the study and the potential for changes even after the start of data collection, our goal was to have all systems easily configurable so updates will be applied without interruption of data collection and in a timely manner. Significant changes were made to a GSCS and IFMS which are described below.

3 Application ImportBlaise2SQL

To accomplish our goals to simplify maintenance of so many Blaise instruments and to be able to post up-to-date collected data on the web portal, we decided to create a table for each Blaise instrument in SQL Server with a subset of fields. We considered the possibility of using the Blaise Datalink Component to create and update these tables. Although it looked like an easy solution to our needs, we soon realized that maintenance of a .BOI file for each Blaise instrument could be time consuming and would require extensive support from an experienced Blaise programmer because the list of fields to pass from the Blaise database to SQL table was not stable. We decided to develop a .Net application that will use a special MS SQL table to create all needed MS SQL tables on the fly and populate them with the most recent data from the Blaise databases.

Development of this ImportBlaise2SQL application began with the creation of a detailed list of Blaise databases and variables which are required for tracking, event coding, preloads for follow-up instruments, Data Entry in Control System and web portal. These variables are stored in an SQL table

(Tbl_Config) to allow the application to import data from these variables into the Blaise databases through the nightly job.

Each record in the Tbl_Config table consists of:

- Original Blaise database name
- Fully qualified variable name along with its size and type in the Blaise database
- Variable column name and its expected size and type in the target SQL table
- Status of the variable to process: “Active” or “Not Active”

The program uses the Blaise Component Pack (BCP) to open the Blaise database to access data collected in the Blaise interview. It then reads the variables given in the configuration table from the appropriate Blaise database and exports them into temporary tables in a SQL Server database as shown in the example below.

Figure 3: Example of Creating Temporary Table in SQL Server

```
//Open Blaise Database
BlAPI3A.Database db = dbMgr.OpenDatabase(DBPath);
db.AccessMode = BlAccessMode.blamShared;
db.Connected = true;
...
//dynamically create columns for temporary table
SQLCols = SQLCols + "[" + (string)dtrCat2["ColName"] + "],";
SQLVars = SQLVars + "@" + string)dtrCat2["ColName"].ToString().Replace(".", "_") + ",";
command.Parameters.AddWithValue("@" + (string)dtrCat2["ColName"].ToString().Replace(".",
"_"), FormatData(db.get_Field((string)dtrCat2["VarName"])));
...
command.CommandType = CommandType.Text;
command.CommandText = "Insert into TEMP_" + DBName + "(" + SQLCols + ") values " + "(" +
SQLVars + ")";
command.ExecuteNonQuery();
...
```

The variables and databases from Blaise are split into temporary tables, and are populated each night by a scheduled job.

After the completion of data importing steps, this program invokes an SQL stored procedure which reads the variables from the temporary tables to identify the newly completed or updated cases and computes their corresponding status codes to update the Control System. Based on these variables, the stored procedure also determines the records of DBS, Daily Diary and Saliva, and Actigraphy data which are then made available on a secure web portal for client viewing and data entry purposes.

For the baseline of WFHS we created and processed six temporary tables with about one hundred variables. When new CAPI instruments were added for the next wave, adding new tables for them in SQL Server was as simple as inserting rows into the Tbl_Config. During the period when all CAPI and CATI interviews were in progress for all four waves, twenty-two tables were updated every night, and data from more than eight hundred variables was read from the Blaise databases and written to SQL Server tables. When a wave of data collection was completed, the status in Tbl_Config was changed to “Not Active” for its corresponding Blaise variables, and the tables would then not be updated by the nightly process. The data from completed interviews is, however, still available to be used for follow-up interviews.

4 Modifications to IFMS

IFMS is used for almost all field studies conducted by RTI International and is a web application responsible for electronic assignment and transfer of cases to field staff, standard case status reports, data transmission, field monitoring, interviewer production, and laptop case management. IFMS is set up for each study and standard reports usually cover mostly all aspects of data collection.

4.1 Custom Reports

The participants in the study in question represent two disparate workplace cohorts – a white-collar, high-tech industry and a long-term care industry. Approximately 600 workplace supervisors, 3,000 employees, 1,200 spouses or partners and 1,200 children (between the ages of 9 and 17) of participating employees are recruited into the study during the baseline period. Data is collected from distinct work units on a rolling schedule across time per industry. Due to the complexity of the study, we developed different types of customized reports to oversee the data collection efforts. Below are just few examples:

- Interview Status Reports by Industry/Worksite
- Basic Health Measures Reports
- Consent Reports by Industry
- DBS/Actigraphy Reports by Industry/ Worksite

In addition, in order to monitor data collection per industry and site in each wave, a dashboard report was developed in .Net which offered a summarized simple overview of the interview counts associated with all instruments in each wave grouped by industry and worksite.

Figure 4: Partial Dashboard Report with Overall Numbers for 4 Waves and Each Site

02/07/2012	Overall		Site 1.0		Site 2.0		Site 3.0		Site 4.0	
	#	%	#	%	#	%	#	%	#	%
Manager	**876**		**14**		**6**		**19**		**34**	
Pending refusal	2	0.23%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Pending other	19		0		0		0		0	
Not eligible	10		0		0		2		0	
Final refusal	33	3.81%	0	0.00%	0	0.00%	1	5.88%	2	5.88%
Final other	37		0		0		0		1	
Completes	775	89.49%	14	100.00%	6	100.00%	16	94.12%	31	91.18%
Employee	**3534**		**82**		**59**		**93**		**134**	
Pending refusal	3	0.08%	0	0.00%	0	0.00%	0	0.00%	0	0.00%
Pending other	80		0		0		0		0	
Not eligible	79		0		2		1		6	
Final refusal	338	9.78%	10	12.20%	11	19.30%	23	25.00%	20	15.63%
Final other	204		3		1		4		12	
Completes	2830	81.91%	69	84.15%	45	78.95%	65	70.65%	96	75.00%

4.2 Changes to IFMS website

There are more than hundred reports available in IFMS for WFHS now. To make it possible to easily find a report for review, IFMS website was modified to dynamically create a list of reports for a selected wave only.

To help Field Supervisors to monitor all the paperwork in the field, a new web page was designed in the IFMS website to display all documents collected during the CAPI interview. The list includes consent forms, Daily Diary receipt forms, etc., with up to 15 forms per respondent. The number of forms and the type of forms are specific to a case and is based on the responses to the consent questions in the various Blaise instruments. The SQL tables which are populated with Blaise data using the ImportBlaise2SQL application simplified the task of generating the list of forms that are expected from a respondent. The web page allowed the Field Supervisor to select or type in a Case ID and get the list of forms associated with the case. The Field Supervisor could then verify the case folder and ensure that all the forms pertaining to the case had been received.

Figure 5: List of Forms Expected for the Case Collected in Three Blaise Instruments

Case ID	Form Name	Form Type	Instrument	Blaise Variable
312MA99E	PINK - D	DBS	Aw12m	DB_12.DB_3
312MA99E	PURPLE - E	Actigraphy	Aw12m	ACT_12.ACT_1
312MA99E	GREEN - C	Employee - HIPAA Waiver	Em12m	LE_12.LE_HIPCK
312MA99E	YELLOW - A	Employee - Workplace CAPI	Em12m	RD_12.RD_CNST1
312MA99E	N/A - Verbal	Spouse	Sp12m	SD_12.SD_CNST4

5 Modifications to Control System

The Control System is where all data collection activity is centrally tracked and monitored. The RTI General Survey Control System provides the database structure and a website that manages the events for:

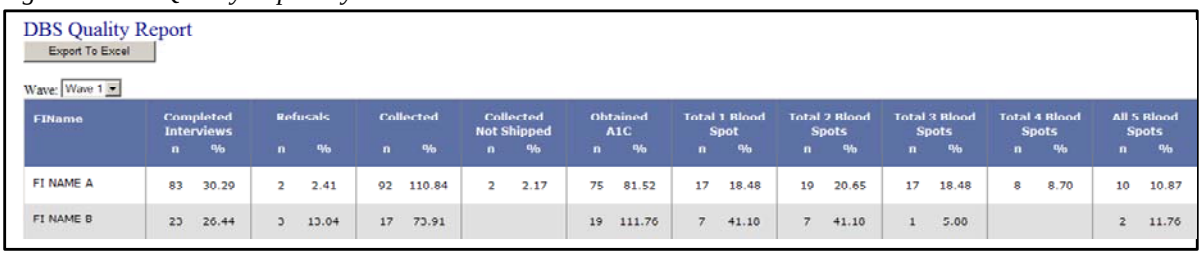
- Completion status for each case in all instruments
- Incentive Mailings
- Interview Verification

While the base structure was sufficient to track the events, the system had to be modified to accomplish specific requirements for WFHS.

5.1 Custom Reports

The reporting system was enhanced to provide 18 custom reports to accommodate the study needs of monitoring the biomarkers, Daily Diary and Saliva eligibility status. Each report is provided by wave and the data could be exported out as an Excel spreadsheet for easier analysis. The following screen displays one such report developed to monitor the quality of the DBS collected for each interviewer.

Figure 6: DBS Quality Report by Field Interviewer


DBS Quality Report

Export To Excel

Wave: Wave 1

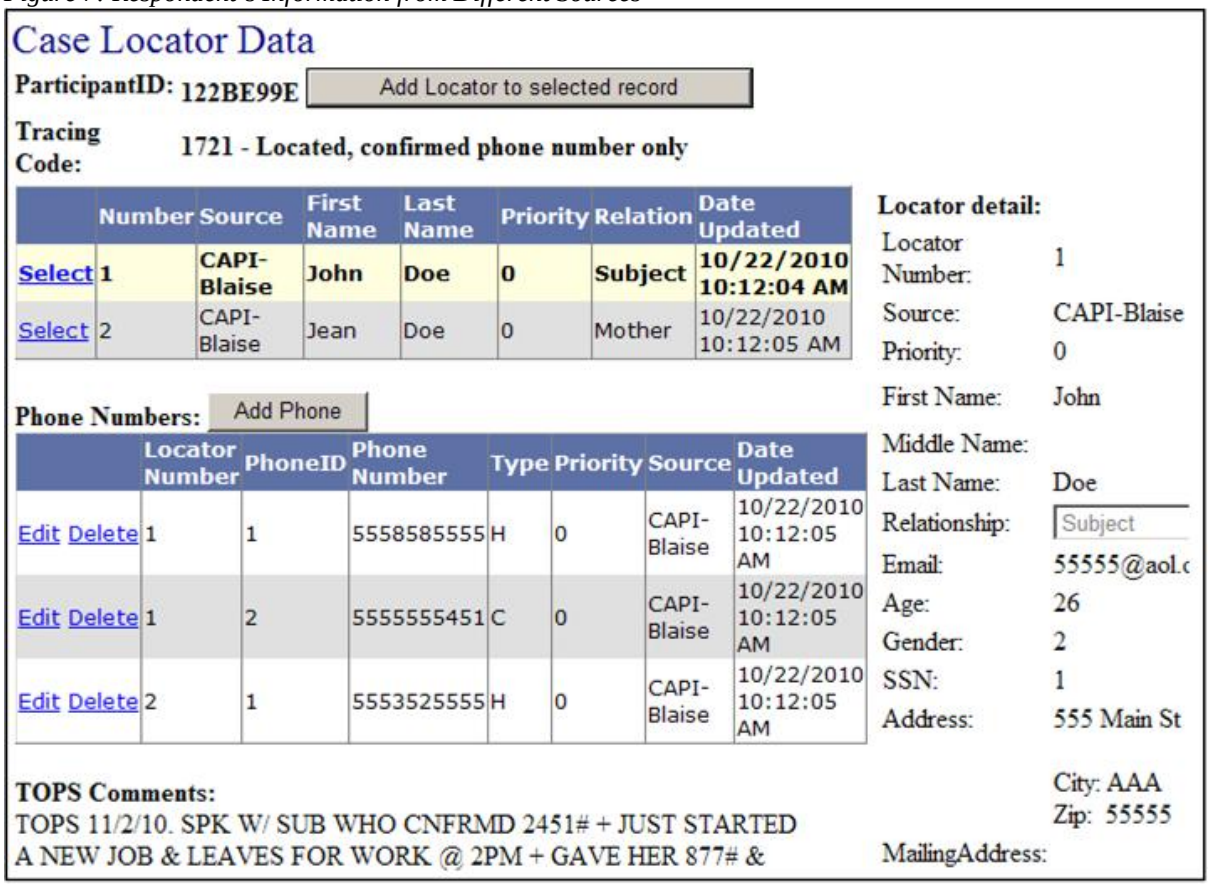
FIName	Completed Interviews n	Completed Interviews %	Refusals n	Refusals %	Collected n	Collected %	Collected Not Shipped n	Collected Not Shipped %	Obtained A1C n	Obtained A1C %	Total 1 Blood Spot n	Total 1 Blood Spot %	Total 2 Blood Spots n	Total 2 Blood Spots %	Total 3 Blood Spots n	Total 3 Blood Spots %	Total 4 Blood Spots n	Total 4 Blood Spots %	All 5 Blood Spots n	All 5 Blood Spots %
FI NAME A	83	30.29	2	2.41	92	110.84	2	2.17	75	81.52	17	18.48	19	20.65	17	18.48	8	8.70	10	10.87
FI NAME B	23	26.44	3	13.04	17	73.91			19	111.76	7	41.10	7	41.10	1	5.00			2	11.76

The Control System website is an internal website only accessible by RTI staff, but the reports from the Control System needed to be ported to an external web portal so all researchers participating in WFHS could review them. We developed a.Net program that would call the report URL from the Control System and save the results as an html file to be posted on the external web portal. This program is scheduled to copy the Control System reports to the secure web portal every night. As a result, there is no need to copy over the data just for reporting purposes.

5.2 Locator Database

Respondents who left their companies during the course of the study are referred to as Attriters. Attriters need to be traced so that the data can be collected for the subsequent wave. In order to trace them, we extract the contact information collected in the Blaise instruments into SQL tables using the ImportBlaise2SQL application and then move necessary data to the locator tables in the Control System. This latest contact information is exported from the Control System and passed over to the RTI Tracing Operations System (TOPS). The confirmed located information is then brought back into the Control System, so that mail-outs can be sent to the correct address and Attriter Blaise interview could be conducted over the phone.

Figure 7: Respondent's Information from Different Sources



When the CATI Attriter interview is completed, new and updated information about the respondent is used for the next Attriter interview so that interview data for all four waves will be available for researchers at the end of the data collection.

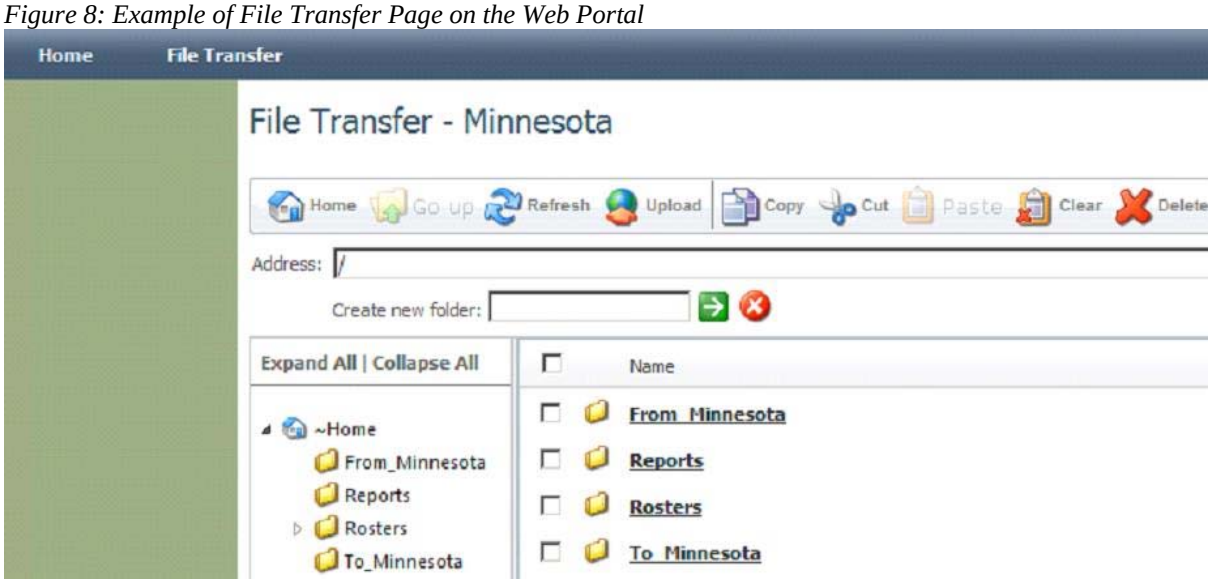
6 Web Portal

The WFHS web portal developed by RTI provides an environment that allows all participating researchers to do file and data exchanges. In addition, Harvard and Penn State have data entry web pages to record results of DBS, Actigraphy, Saliva, and Daily Diary data collections. The portal also contains various status reports to inform researchers about the status and outcomes from the CAPI interviews.

6.1 Web Portal User Access

Because the participating parties have different needs, the web portal requires user authentication for data security purposes. Access to different web portal features is controlled through roles assigned to menu items. Each menu item has certain roles assigned to it, such as "Harvard user" or "Minnesota user". When a user from University of Minnesota logs in, they see the menu specific to their user role and permissions. Thus, the role-based menu can be used to hide information based on the person's role in the study. Only RTI International, being the Data Coordinating Center for this study, has links between participants' answers in CAPI interviews and all other components of data collections. Even reports are specific to a user and only top level reports are available to everybody.

Figure 8: Example of File Transfer Page on the Web Portal



6.2 Web Portal Data Entry

As described above, when Blaise data is extracted for completed CAPI interviews into Control System tables by the ImportBlaise2SQL application, SQL server tables with fields for DBS, Actigraphy, and Daily Diary are updated and posted to an external SQL server database for processing. A "Harvard user" or "Penn user" enters laboratory results and status of their data collection into the web portal.

Another scheduled job picks up updates from the external web portal and uploads them back into Control System. At that time new statuses are applied and reports show up-to-date statuses from all sources.

7 Support and Development of Blaise Instruments

In previous sections we described some of the new development that was required for WFHS. We also used many existing tools and applications that helped us collect data efficiently and successfully.

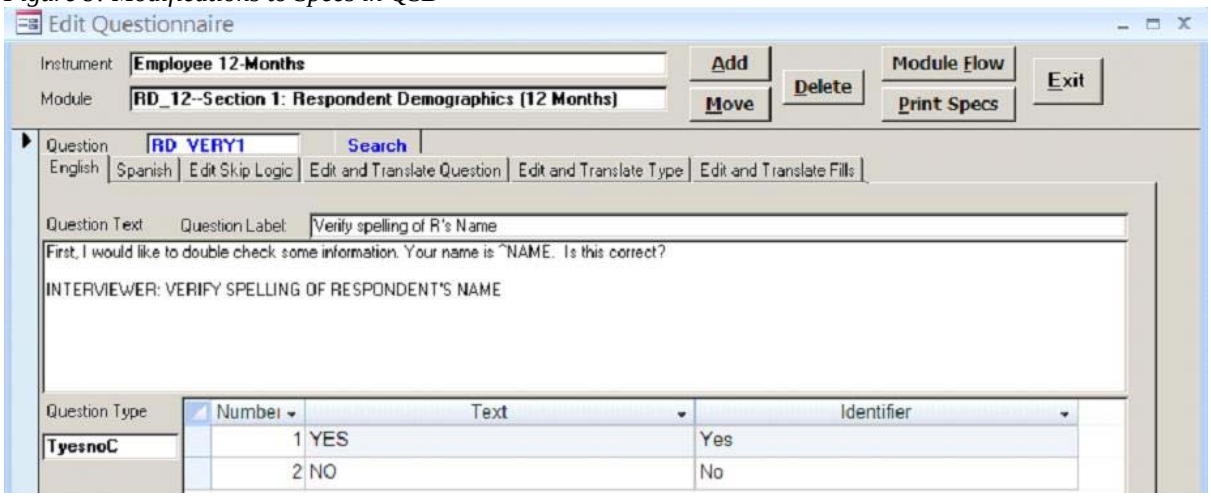
7.1 Use of Questionnaire Specification Database (QSD) to Develop Instruments

Given a very tight schedule to develop Blaise instruments, especially for the follow-up interviews that happened every 6 months, we made extensive use of our Questionnaire Specification Database (QSD) for developing all CAPI/CATI Blaise instruments. QSD consists of tiers for the user interface and business rules, and a back-end Microsoft Access database. It helps spec writers and Blaise developers streamline the process of creating Blaise instruments. A key feature of QSD is its ability to track all changes made to the specifications. This is very helpful for programmers and translators to check for updates after revisions have been made by the spec writer.

Since WFHS is a large longitudinal study utilizing 20 CAPI and 2 CATI Blaise instruments among four waves of data collection, it was and still is a challenging task to maintain and manage all facets of the instruments development process. Although many questions in follow-up interviews are repeated, text for them could be changed and new questions could be added. So for all follow-up interviews new specs were originally to be created by spec writers and then implemented by the Blaise programmers. Instead, to simplify the specification process and to quickly prepare new Blaise instruments, we used QSD to add new modules by duplicating existing modules from previous waves and to allow spec writers to enter modifications directly in QSD. Every time a change was made, information was recorded about who made the change, the date and time, and details outlining the modifications made. After spec writers were finished with the changes, programmers used QSD to

generate follow-up Blaise instruments and produce specification documents that are complete with question text, response options, fills, and routing comments. Since specifications and Blaise code are both generated from within QSD, both documentation and Blaise code are synchronized and linked throughout the development process.

Figure 9: Modifications to Specs in QSD



7.2 Updates of Blaise Instruments on Field Interviewer Laptops

When data collection for WFHS started in 2009, only 6 baseline Blaise instruments were on FI laptops. Since then all updates to existing instruments and additions of later instruments were done remotely using IFMS without interruption to data collection.

During this time, the economic situation of companies participating in the study was changing and new questions needed to be added to the instruments to reflect those changes. But for modifications that alter the structure of a Blaise data model, it is difficult to do: any existing production Blaise data files must be upgraded along with the new data model files. Although the Manipula setup for this procedure is very simple, many steps need to be taken to avoid loss of data: creation of backup folders, copying of a number of files, the preparation of Manipula setups for the upgrade and so forth. It can be an extremely complex process, especially when instruments reside on multiple computers, possibly with different versions.

This complicated task is simplified through the use of a special program, Upgrade Blaise Instrument (UBI), developed some time ago at RTI International. UBI utilizes the version number which exists for each Blaise project, data model, and data file. When a new version of the Blaise instrument is ready for the field, UBI is used to upgrade the combined CAPI database that resides at RTI and to create a package of files that are required to upgrade FI laptops. The package then is sent via transmission and executed on a laptop without FI interaction. Results of the upgrade are automatically sent back to RTI so we can monitor the status of the upgrade and the version number of the instrument. On average, every instrument in the field was upgraded twice.

8 Conclusion

Data collection for WFHS has continued successfully for more than two years, and two waves are already completed. The third and fourth are scheduled to be completed by the end of this year.

In the process of supporting so many Blaise instruments at once, the following are some of the challenges we encountered and overcame:

- Data collected in CAPI should be available to participating researchers as soon as possible on the RTI web portal.

- Reports should be split up by key variables (industry/site/etc) and on the contrary combined for all waves
- Changes to the instruments can require upgrades to Blaise data files that should very quickly be applied to FI laptops in production
- Data from different components of data collection should be merged for a respondent without disclosure of the respondent's identity
- All respondents who completed baseline interview should be tracked and data for subsequent waves should be collected - even if a respondent no longer worked at a participating company

Of course there were and still are other challenges inherent in supporting Blaise instruments for this study that are not part of this paper. The applications and systems described here addressed the requirements of the study. They help make data collection for WFHS efficient and will provide all the participating parties with accurate data for their research.

9 References

Bray, J., Almeida, D., Buxton, O., Dearing, J., Kelly, E., King, R. "An integrative, multi-level, and multi-disciplinary research approach to challenges of work, family, and health." 2012

Lilia Filippenko, Valentina Grouverman, Joseph Nofziger, "Questionnaire Specification Database for Blaise Surveys", 2007, Proceedings of 11th International Blaise Users Conference

Lilia Filippenko, Joseph Nofziger, and Gilbert Rodriguez, "Automatic Upgrade of Production Blaise Instruments", 2003, Proceedings of 8th International Blaise Users Conference

10 Acknowledgement

This research was conducted as part of the Work, Family and Health Network (www.WorkFamilyHealthNetwork.org), which is funded by a cooperative agreement through the National Institutes of Health and the Centers for Disease Control and Prevention: Eunice Kennedy Shriver National Institute of Child Health and Human Development (Grant # U01HD051217, U01HD051218, U01HD051256, U01HD051276), National Institute on Aging (Grant # U01AG027669), Office of Behavioral and Science Sciences Research, and National Institute for Occupational Safety and Health (Grant # U01OH008788, U01HD059773). Grants from the William T. Grant Foundation, Alfred P Sloan Foundation, and the Administration for Children and Families have provided additional funding. The contents of this publication are solely the responsibility of the authors and do not necessarily represent the official views of these institutes and offices. Special acknowledgement goes to Extramural Staff Science Collaborator, Rosalind Berkowitz King, Ph.D. and Lynne Casper, Ph.D. for design of the original Workplace, Family, Health and Well-Being Network Initiative.