METHODS REPORT

New Technologies and Tools for Study Management: Lessons Learned from Designing and Implementing a Web-Based Data Management System for a Multisite Longitudinal Intervention Study

Lauren Courtney, Emily Warmoth, Margaret Rodan, Kathy Katz, Siva Subramanian, and Michele Kiely

December 2008

RTI Press

RTI INTERNATIONAL

About the Author

Lauren Courtney, MPH, is a research epidemiologist at RTI International in Washington, DC.

Emily Warmoth, MSIS, is a research programmer/analyst at RTI International in Research Triangle Park, NC.

Margaret Rodan, ScD, is the GirlTalk project manager at Georgetown University.

Kathy Katz, PhD, is a clinical psychologist for GirlTalk at Georgetown University.

Siva Subramanian, MD, is the principle investigator for GirlTalk at Georgetown University.

Michele Kiely, DrPH, serves as the project officer of the DC Initiative for the Eunice Kennedy Shriver National Institute of Child Health and Human Development.

RTI Press publication MR-0005-0812

Suggested Citation

Courtney, L., Warmoth, E., Rodan, M., Katz, K., Subramanian, S., and Kiely, M. (2008). New Technologies and Tools for Study Management: Lessons Learned from Designing and Implementing a Web-Based Data Management System for a Multisite Longitudinal Intervention Study. RTI Press publication No. MR-0005-0812. Research Triangle Park, NC: RTI International. Retrieved [date] from http://www.rti.org/rtipress.

This publication is part of the RTI Press Methods Report series.

RTI International
3040 Cornwallis Road
PO Box 12194
Research Triangle Park, NC
27709-2194 USA

Tel: +1.919.541.6000
Fax: +1.919.541.5985
E-mail: rtipress@rti.org
Web site: www.rti.org

doi:10.3768/rtipress.2008.mr.0005.0812

www.rti.org/rtipress

New Technologies and Tools for Study Management: Lessons Learned from Designing and Implementing a Web-Based Data Management System for a Multisite Longitudinal Intervention Study

Lauren Courtney, Emily Warmoth, Margaret Rodan, Kathy Katz, Siva Subramanian, and Michele Kiely

Abstract

In epidemiologic research, effective tools for study management improve data quality, staff/time efficiency, and cost-effectiveness and strengthen study outcomes. The Web-based data management system (DMS) is a relatively new technology in epidemiologic research which provides significant benefits of these kinds. This paper addresses the design, technical implementation, and field experiences of using a Web-based DMS for the *GirlTalk for Teen Moms Study*, conducted as part of the NIH-DC Initiative to Reduce Infant Mortality in Minority Populations in Washington, DC (the NIH-DC Initiative).

RTI developed a Web-based DMS for the GirlTalk study to assist field staff from multiple sites in efficiently managing participant activities and data collection for this multiple component, multisite longitudinal study. The GirlTalk DMS monitors all activities for screening, recruitment, enrollment, in-depth interviews, randomization, interventions, repeated measures, outcomes, adverse events, and follow-up activities for each participant over the course of her 2-year enrollment in the study. Field staff from multiple sites use the DMS to manage more than 100 possible activities with a complex series of dependencies between activities that lead from one activity to the next.

We compare the features of the Web-based DMS used for GirlTalk with those of a personal computer (PC)-based DMS used previously by this research group. Our comparison presents the efficiencies obtained, including real-time data access, simplified software management, data security, and anytime/anywhere access for all users. We outline cost considerations for the Web-based DMS versus the PC-based system in terms of study size, study duration, number of sites, and technical infrastructure. We illustrate several features of the DMS, including the role-based menu, events tracking, appointment scheduling, shared calendars, report generation, access to study documents, and uploading of data files to RTI International secure servers. Finally, we compare the value of the Web-based tool versus the PC-based tool for study management.

Contents

RTI INTERNATIONAL

Introduction

Researchers require effective study management tools to improve data quality, staff/time efficiency, and cost-effectiveness, and to strengthen study outcomes (Unutzer et al., 2002; Winget et al., 2005). The Web-based data management system (DMS) is a relatively new technology in epidemiologic research that provides significant benefits in these areas. Access to the Web is becoming widely available, with more than 70 percent of Americans reporting Internet access in December 2007 (Miniwatts Marketing Group, 2008).

The remainder of this report discusses the methodology used to design and implement a Web-based DMS for a multiple component, multisite longitudinal study. We compare features of the Web-based DMS that we are currently using with those of a personal computer (PC)-based DMS that we used in our earlier study. This comparison will assist the reader in assessing the relative value of the Web-based tool versus the PC-based tool.

The NIH-DC Initiative to Reduce Infant Mortality

We have conducted two studies—the *GirlTalk for Teen Moms Study* and *Interventions for Risk Factors in Pregnant Women in Washington, DC: An Integrated Approach* (Project DC HOPE)—as part of the NIH-DC Initiative to Reduce Infant Mortality in Minority Populations in Washington, DC (NIH-DC Initiative). Washington, DC, has historically reported some of the highest infant mortality rates in the United States, with 11.0 infant deaths per 1,000 live births in 2004, as compared with a national average of 6.6 per 1,000 (Munson and Sutton, 2006). With particularly high rates, African Americans in Washington, DC, were reported to have 14.5 infant deaths per 1,000 live births in 2002 (Kochanek et al., 2004).

In response to these high mortality rates, beginning in 1993 the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD) provided 5-year grants to collaborating universities and research institutions in the Washington, DC, area to evaluate multiple contributing factors to infant mortality and develop effective community-based intervention strategies. The NIH-DC Initiative is currently in its third phase (conducted 2004–2009); it comprises GirlTalk and four other NIH-DC Initiative studies. We conducted Project DC HOPE during Phase 2 (1998–2004); it had study activities and data management requirements similar to those of GirlTalk. (See Appendix A for descriptions of the GirlTalk and DC HOPE projects.)

Previous studies in the NIH-DC Initiative, including DC HOPE, used PC-based data management systems requiring nightly data transmissions to and from each site to synchronize study data. In developing study management tools for Phase 3, RTI staff reviewed the DMS used in Phase 2 and developed an innovative DMS based on current-generation software applications, primarily to improve efficiency and maximize data quality. Specifically, for GirlTalk, we developed a Web-based DMS to assist field staff from multiple sites in efficiently managing participant activities and data collection for this complex, 5-year study.

Procedures for Studies

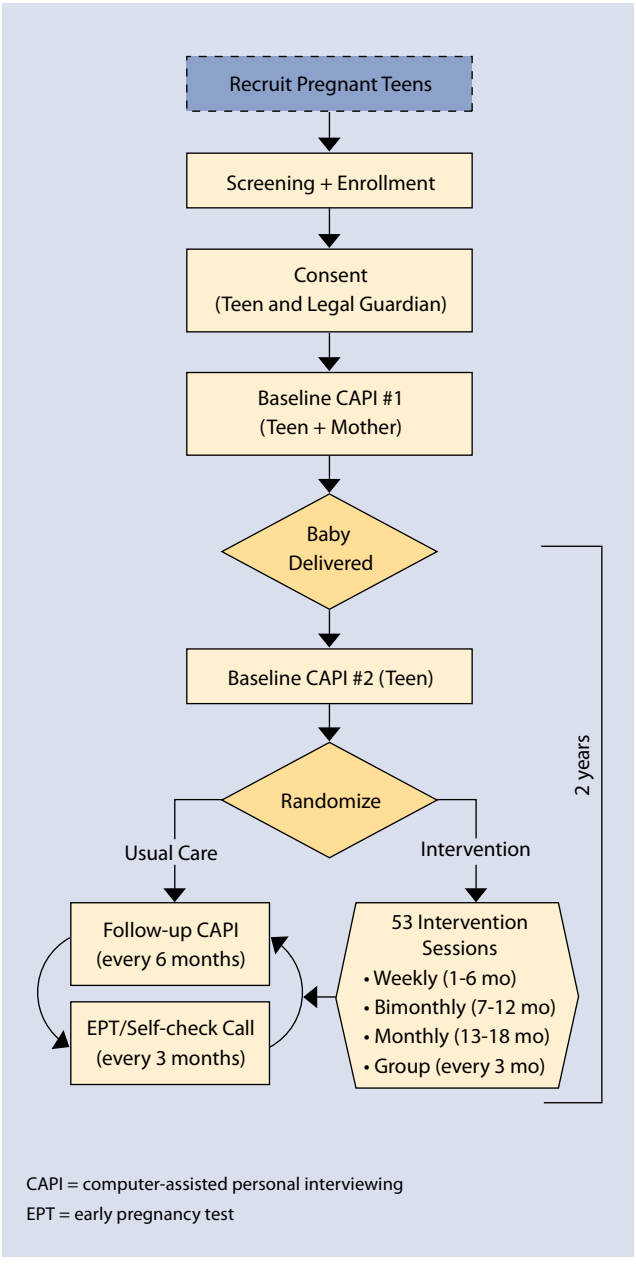
The primary objective of the GirlTalk study, a multisite randomized clinical trial, is to evaluate the effectiveness of a community-based intervention for minority teen mothers with the principal purpose of reducing subsequent teen pregnancies. Various staff from multiple organizations and sites are involved in data collection and implementation of the intervention for this study, each with unique roles and responsibilities. Counselors from Georgetown University (GU) manage participant recruitment, enrollment, baseline interviews, counseling interventions, and overall participant retention for follow-up interviews. Telephone interviewers from Children's National Medical Center (CNMC) conduct baseline and follow-up phone interviews and manage participant retention.

As the NIH-DC Initiative data coordinating center (DCC), RTI staff are responsible for developing the study management system, creating data collection instruments, monitoring data collection processes, and conducting data analysis. During Project DC HOPE, the collaborating research groups and RTI performed the same roles that they are now performing for GirlTalk. In addition, the tasks conducted by staff during DC HOPE were similar

to those being conducted for GirlTalk, including frequent follow-up interviews, intervention sessions, mailings, participant retention, and data quality monitoring through the DMS. For these reasons, we believed an in-depth comparison of the DMSs used by these two studies would be a valuable contribution to the methods literature as research of this type moves increasingly to Web-based tools.

Figure 1 provides a brief overview of the GirlTalk procedures.

Figure 1: Flowchart of the GirlTalk study activities



CAPI = computer-assisted personal interviewing
EPT = early pregnancy test

Appendix B provides a complete list of DMS events. In total, the study aims to recruit 340 teens at staggered intervals and enroll them for a 2-year period. Counselors recruit pregnant teens through in-person visits to and referrals from more than 200 public schools, community centers, and health clinics throughout the Washington, DC, area. After conducting a screening, GU counselors enter information on the eligible participants (i.e., ages 15 to 19 years, African-American or Latina, and primiparous) into the DMS. GU counselors enroll teens and their mothers (or a mother figure) as a dyad. Legal guardian written consent is required for minors (participants under age 18).

CNMC interviewers collect evaluation measures at baseline and 6-month intervals for the teen and at baseline and 12-month intervals for the mother. GU counselors conduct the prenatal baseline interview with the teen and mother via computer-assisted personal interviewing (CAPI) during a home visit. CNMC interviewers conduct the teen's postpartum baseline interview and follow-up interviews via CAPI by telephone. Additionally, they contact the teens every 3 months to verify their pregnancy status through an early pregnancy test or self-report. GU counselors use information in the DMS to create preaddressed mail merge letters and mail a $15 incentive check to each teen or mother who has completed an interview.

The DMS displays an alert when a teen is ready for randomization. GU counselors initiate the randomization, and the DMS randomly selects the group assignment for each enrolled teen—either intervention or usual care. The DMS allows randomization only after CNMC interviewers upload the baseline evaluation data and add participant delivery data to the DMS. Counselors give teens in the intervention group cell phones and conduct counseling sessions by telephone to promote behaviors that may postpone another pregnancy. Beginning the interventions at 4 to 8 weeks postpartum, counselors conduct counseling weekly for the first 6 months, biweekly for the next 6 months, and monthly for the next 6 months; they provide no counseling for the remaining few months through 2 years. Overall, counselors and interviewers

are responsible for managing more than 100 possible activities for each participant, with a complex series of triggers that lead from one activity to the next.

Design of the Data Management System

To develop the GirlTalk DMS, GU and CNMC staff worked collaboratively with the RTI data management team to define the design goals based on how the DMS would be used. RTI used the protocol to complete the design process and documented the program requirements.

Design Goals for the DMS

In planning the design of the DMS for the GirlTalk study, RTI and GU reviewed the PC-based DMS developed for Project DC HOPE and developed the following list of goals:

- **Improve data quality:** Because multiple users enter and modify DMS data, quality assurance/quality control (QA/QC) features such as data verification and validation checks are needed to ensure that clean data are entered into the DMS. QA/QC checks are in place on the Web-based electronic data entry forms and events.
- **Increase protocol adherence:** During the 2-year study, each of the 340 participants must complete more than 100 activities, including data collection and intervention sessions. Through generating reports, triggering subsequent events, and providing reminders, the DMS assists staff in monitoring the events for each participant and ensuring that study activities are completed in compliance with the protocol.
- **Provide real-time data:** The GirlTalk DMS has real-time data, updated continuously and available at any location to link the multiple study sites effectively.
- **Increase efficiency/responsiveness:** Allowing multiple users to enter and access data in real time reduces the need for telephone calls and e-mail communications between sites about activities that affect their responsibilities.
- **Reduce staff time/burden:** Staff complete all data forms electronically so that forms with common data elements, such as contact information, can be linked and prepopulated, which eliminates the need to enter the same data in multiple places.
- **Increase participant retention:** DMS reports allow study investigators to monitor participant progress and alert them about difficult-to-reach participants so that study staff can implement appropriate retention strategies before participants become lost to follow-up.

The Design Process

The RTI team began development of the DMS by creating a detailed flow chart of the activities required for each participant throughout the 2-year enrollment period. The structure of the DMS was based on the three main components of the study: recruitment, intervention, and evaluation. We determined the following elements of each component collaboratively:

- Whether staff should complete a hard copy data form or an electronic one.
- Whether staff need a report to track a specific activity for purposes of scheduling or to monitor progress of the activity for purposes of quality control, and if so, what key data staff need to meet the purpose of the report.
- The window of time allocated to complete each activity. We used this information to develop reports that would alert GirlTalk study staff to upcoming activities so they could commit the necessary resources for completing the activity within the appropriate time frame.
- Individuals' level of access to specific data within the DMS, based on their roles on the study. For example, to maintain the integrity of the study design, interview staff conducting the evaluation needed to be blinded from the randomization result.

The RTI developers used the DMS design specifications and requirements documents to plan for and implement the programming and testing.

Technical Implementation of the Web-Based DMS

Description of PC-Based and Web-Based Systems

Figure 2 depicts the PC-based system that we used in the earlier DC HOPE project (see Appendix A). In this system, a central server at the DCC initiated all data transmissions. Each night the central server dialed into the systems at each site and transferred the data. This design required successful nightly data transmissions to and from each site via point-to-point modem transmission. RTI technical staff had to ensure that the central server PC was placed in a secure physical location.

Figure 3 shows the Web-based system that we are using for the GirlTalk study. Computers at GU and CNMC access the DMS through the Web. Field staff upload audio computer-assisted self-interviewing (ACASI) and CAPI data using the DMS. ACASI and CAPI records that remain on the laptops after uploads have no identifiable data. GirlTalk staff members usually connect over their university or hospital local area network (LAN). Occasionally users will connect from home using a dial-up connection.

Figure 2: PC-based system for the DC HOPE project

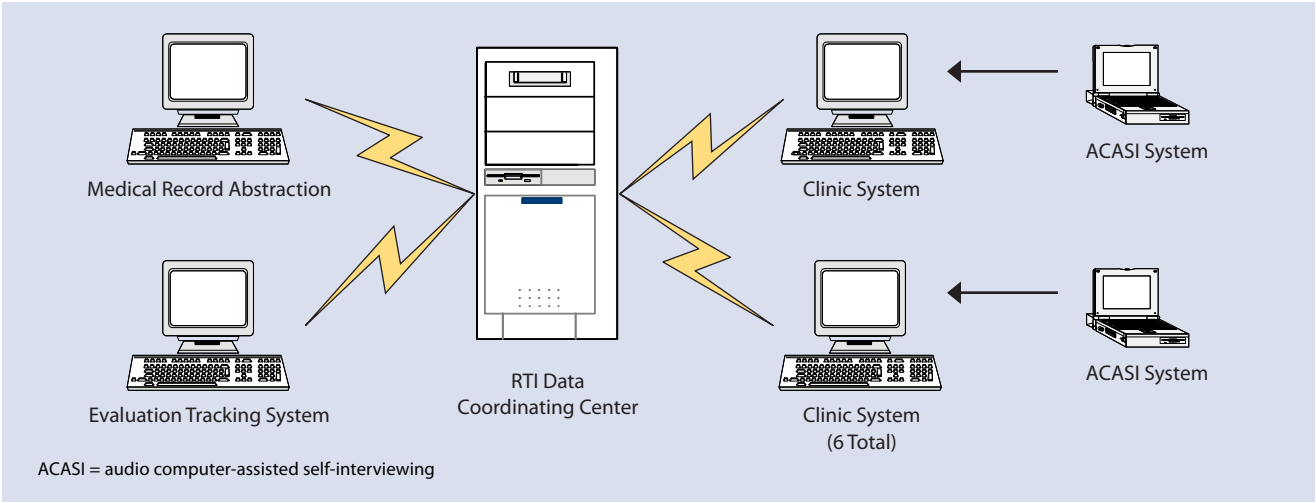
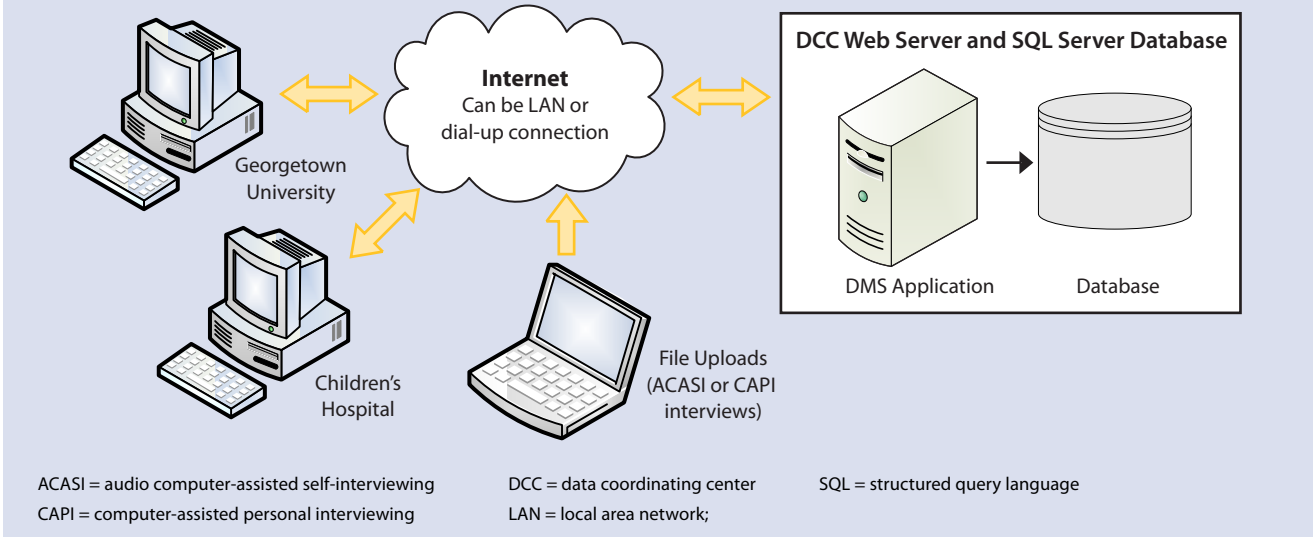


Figure 3: Web-based system for the GirlTalk study



The Web-based model eliminates the problems of delayed data transmission. We no longer need a central server PC to manage the nightly data transmission. RTI stores data in a secure Microsoft SQL Server database. Because RTI stores all DMS data in a central database, data are available to users as soon as they are entered in the DMS or after interview data from a laptop are uploaded. No additional synchronization of the data is needed.

Comparison of the DMS Systems

Table 1 compares the advantages and disadvantages of the PC- and Web-based systems. These comparisons guided the decision to use a Web-based system to increase the overall efficiency of the management of the study. Although programmers faced a learning curve and we had higher initial expenses in creating a Web-based DMS, overall the Web-based system was able to improve the access, security, monitoring, and sharing of data by all GirlTalk project staff.

Table 1a: Comparison of the DMS systems: advantages of a PC-based system

Feature	PC-Based System (DC HOPE)	Web-Based System (GirlTalk)
Offline availability	Internet connection was not needed to access and use the DMS.	Internet connection is required to access and use the DMS.
Learning curve for programmers	The application development tool (Microsoft Access) and the Windows environment were familiar to the programming team at the time (2000).	Developers needed to learn the chosen Web development technology (ASP.NET version 1.1). However, building the GirlTalk DMS with widely used current-generation tools (.NET and SQL Server) will extend the useful life of the system.
Richness of user interface	Applications that run on PCs (e.g., Microsoft Word or Outlook) have a rich and "smooth" interface, allowing users to navigate through many tasks on a complex screen. Because the application runs locally, responsiveness is quick.	For Web-based applications, the user interface toolset is smaller and more limited. With Web pages the user typically performs one simple task per screen. Response time can be slower. When a user performs an action on a Web page, such as clicking a button, it often involves a trip to the server and back.
Costs for initial development	Lower costs up front for initial development since much of the code could be reused from existing programs at that time (2000).	Higher costs up front for application development. However, once a Web-based system is developed for one study, it may be possible to reuse it for other studies.

Table 1b: Comparison of the DMS systems: advantages of a Web-based system

Feature	PC-Based System (DC HOPE)	Web-Based System (GirlTalk)
Software updates	Whenever the software was modified to fix bugs or add new features, updates had to be transmitted to each PC and installed. Onsite technical support had to be provided for locally installed software.	Updating the DMS software is simple because the application files are housed and managed at the DCC. When the application must be changed, the programmer updates the files on the DCC's Web server. Users immediately see the changes in their Web browser. Not needing to update software installed on individual users' computers speeds the correction of bugs and implementation of changes in study procedures.
DMS availability	The DMS could only be used on specific PCs on which the software had been installed.	Users can access the DMS from any computer that has an Internet connection and the Internet Explorer browser.
Synchronization of data	Data between the sites and the DCC were synchronized nightly, resulting in a delay of at least 1 day.	Data are available as soon as they are entered in the DMS, regardless of time of day.
Physical security of data	All study data, including participant identifiers, were stored on PCs at each site. Thus, the physical security of the PCs was a concern.	Users enter data directly into the Web application, and data are stored in a secure database at the DCC. Data being transferred are encrypted using Secure Socket Layer technology. Only ACASI and CAPI files are stored on PCs, and these files contain no identifying information.
Quality of Internet connection	Users connected via telephone connections, which were unreliable and slow (typically with connection speeds of less than 56 Kbps).	Users connect to the DMS using their university's or hospital's LAN; these tend to be fast, stable, and reliable.
Timeliness of data entry	Users often did not enter data promptly because the DCC could not view the data until a transmission took place.	Because data can be entered at a user's location at any time, users tend to enter data soon after an event has occurred. This improves the timeliness of data and reduces errors.
Data monitoring, sharing, and coordinating	Data were not shared between sites. Data were received and aggregated at the DCC, where monitoring took place after a delay.	Data monitoring, sharing, and coordinating between sites is easy and quick. Study investigators can log in to the DMS and monitor the study in real time.
User training	Training sessions had to be conducted onsite. Users had to be taught the functionality of the DMS screens.	The sites already have Internet connections, so training can be done remotely. Also, many users are familiar with Web browsers, which may reduce training requirements.
Hardware and support	Studies often purchase identical equipment for staff to simplify the support of hardware and software. Studies must consider whether purchased equipment fits within the limitations of the particular computing environment at the site.	With a Web-based system, staff may use their own computers and receive general maintenance and support from their information technology staff. These factors save the project time and money.
Costs for maintenance	Maintenance costs are higher due to travel for setting up and maintaining the PCs, purchasing project-specific hardware and software, and increased hours for technical support.	Overall costs are lower for maintenance because of efficiency and ease of applying software upgrades; in-person site visits are not needed.

ACASI = audio computer-assisted self-interviewing
CAPI = computer-assisted personal interviewing
DCC = data coordinating center
DMS = data management system
LAN = local area network

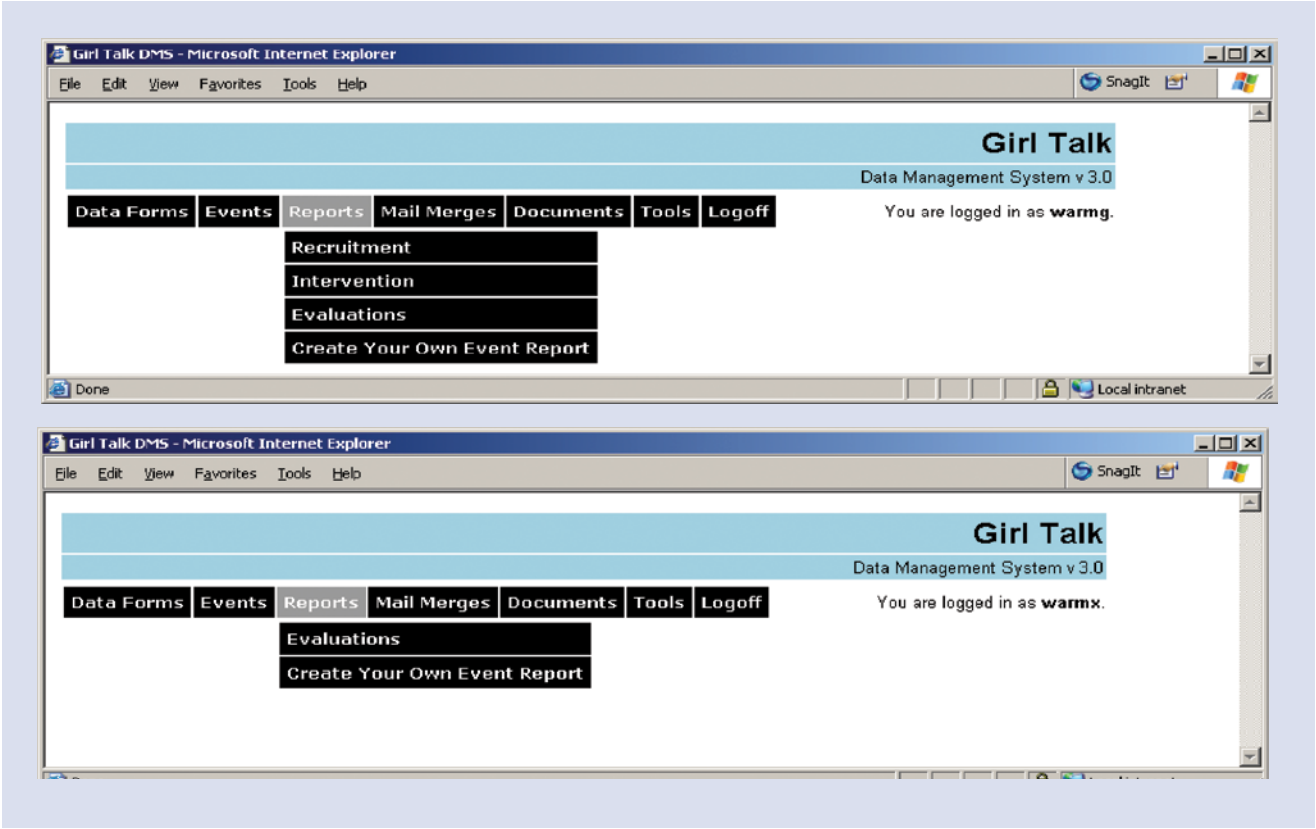
The Web-based DMS contains several unique and essential features that assist GirlTalk counselors, interviewers, and staff in managing the study. Although many of these features can be programmed on both PC-based and Web-based systems, we highlight here the Web-based features.

Role-Based Menu

Access to DMS features is controlled through roles assigned to menu items. Each menu item has certain roles assigned to it, such as the "counselor" role or the "interviewer" role. The advantage is that the DCC can develop and maintain one application for all users instead of needing to maintain the application on multiple laptops and PCs.

Figure 4 shows an image of the different role-based menu options that display for two user types. When the counselors at GU log in, they see the menu shown in the top image in Figure 4. They have access to all of the options under Reports. When the interviewers at CNMC log in, they see the menu in the bottom image, with fewer options under Reports. Thus, managers can use the role-based menu to present

Figure 4: Role-based menu for counselors (top) and interviewers (bottom)



selected information based on a person's role in the study; for example, interview staff conducting the evaluation would be blinded from the randomization result. Programmers can implement this on a PC-based system, but the list of users would need to be updated in a master list and sent out for PC users to update their file.

Events Tracking

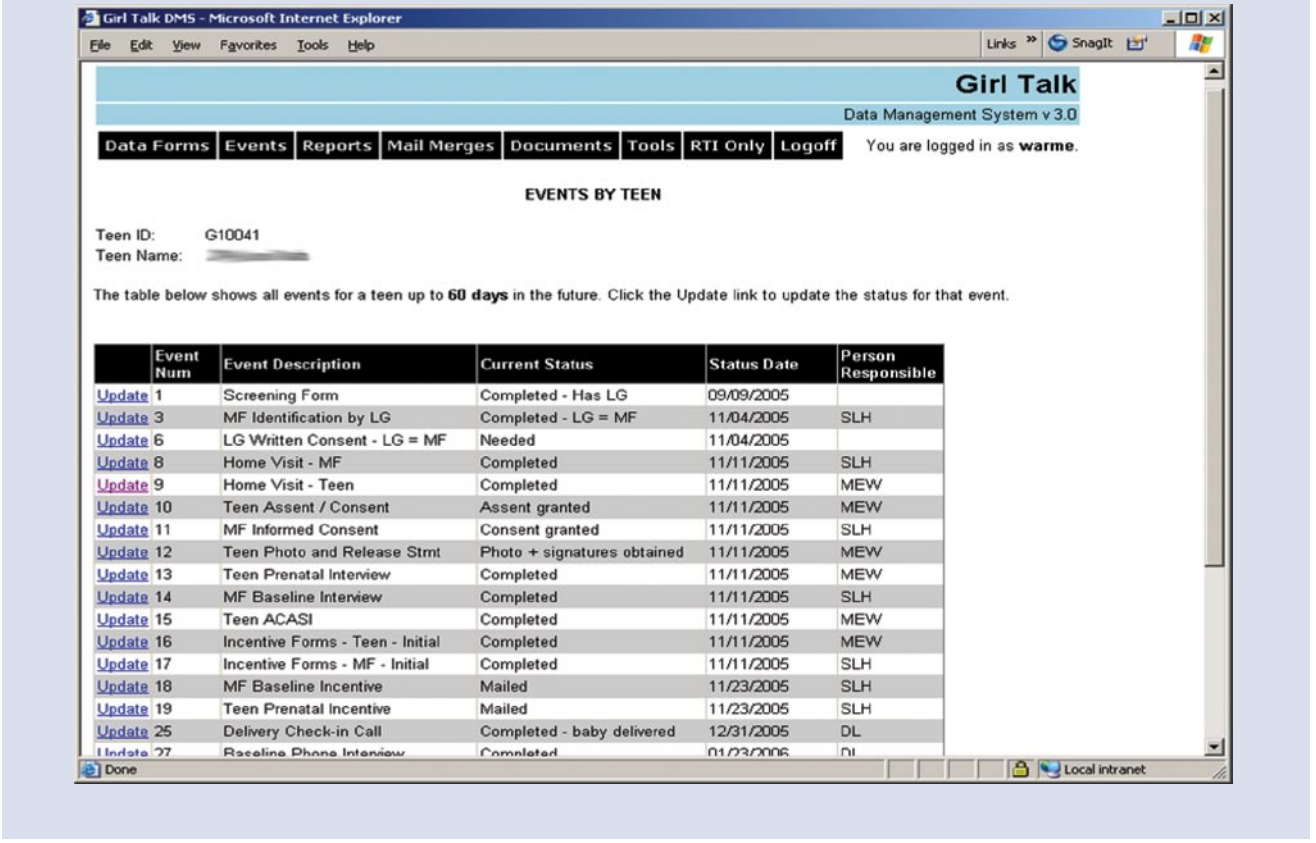
The Events feature allows GirlTalk staff to track the study activities and events for each teen. The Events process is a central piece of the DMS. The Events grid lists a description of the event, the event status, the date the status was updated, and the person responsible for that status.

Figure 5 shows an image of the Events feature. Programmers can implement an Events process on a PC-based system, but staff would not be able to see events from other sites in real time.

All users access the same Events feature, but the DMS determines the subset of events to display based on the user's study site. Therefore, users see only those events that are necessary to perform their jobs.

The Events feature can be thought of as a detailed flow chart for the study. Each event is triggered by a previous event so that the tracking system is as automated as possible. Users are responsible for updating the status of events. The DMS records every activity the user completes as a status change in the Events feature for easy reference and consistency.

Figure 5: Events feature in the Web-based DMS



	Event Num	Event Description	Current Status	Status Date	Person Responsible
Update	1	Screening Form	Completed - Has LG	09/09/2005	
Update	3	MF Identification by LG	Completed - LG = MF	11/04/2005	SLH
Update	6	LG Written Consent - LG = MF	Needed	11/04/2005	
Update	8	Home Visit - MF	Completed	11/11/2005	SLH
Update	9	Home Visit - Teen	Completed	11/11/2005	MEW
Update	10	Teen Assent / Consent	Assent granted	11/11/2005	MEW
Update	11	MF Informed Consent	Consent granted	11/11/2005	SLH
Update	12	Teen Photo and Release Stmt	Photo + signatures obtained	11/11/2005	MEW
Update	13	Teen Prenatal Interview	Completed	11/11/2005	MEW
Update	14	MF Baseline Interview	Completed	11/11/2005	SLH
Update	15	Teen ACASI	Completed	11/11/2005	MEW
Update	16	Incentive Forms - Teen - Initial	Completed	11/11/2005	MEW
Update	17	Incentive Forms - MF - Initial	Completed	11/11/2005	SLH
Update	18	MF Baseline Incentive	Mailed	11/23/2005	SLH
Update	19	Teen Prenatal Incentive	Mailed	11/23/2005	SLH
Update	25	Delivery Check-in Call	Completed - baby delivered	12/31/2005	DL
Update	27	Baseline Phone Interview	Completed	01/23/2006	DL

The Events feature uses a table-driven design in which events and their possible status codes are static values stored in tables in the study database. When a user updates the status of an event, a SQL Server trigger is fired that checks the event tables to determine whether new event(s) need to be created or whether the status needs to be changed on an existing event.

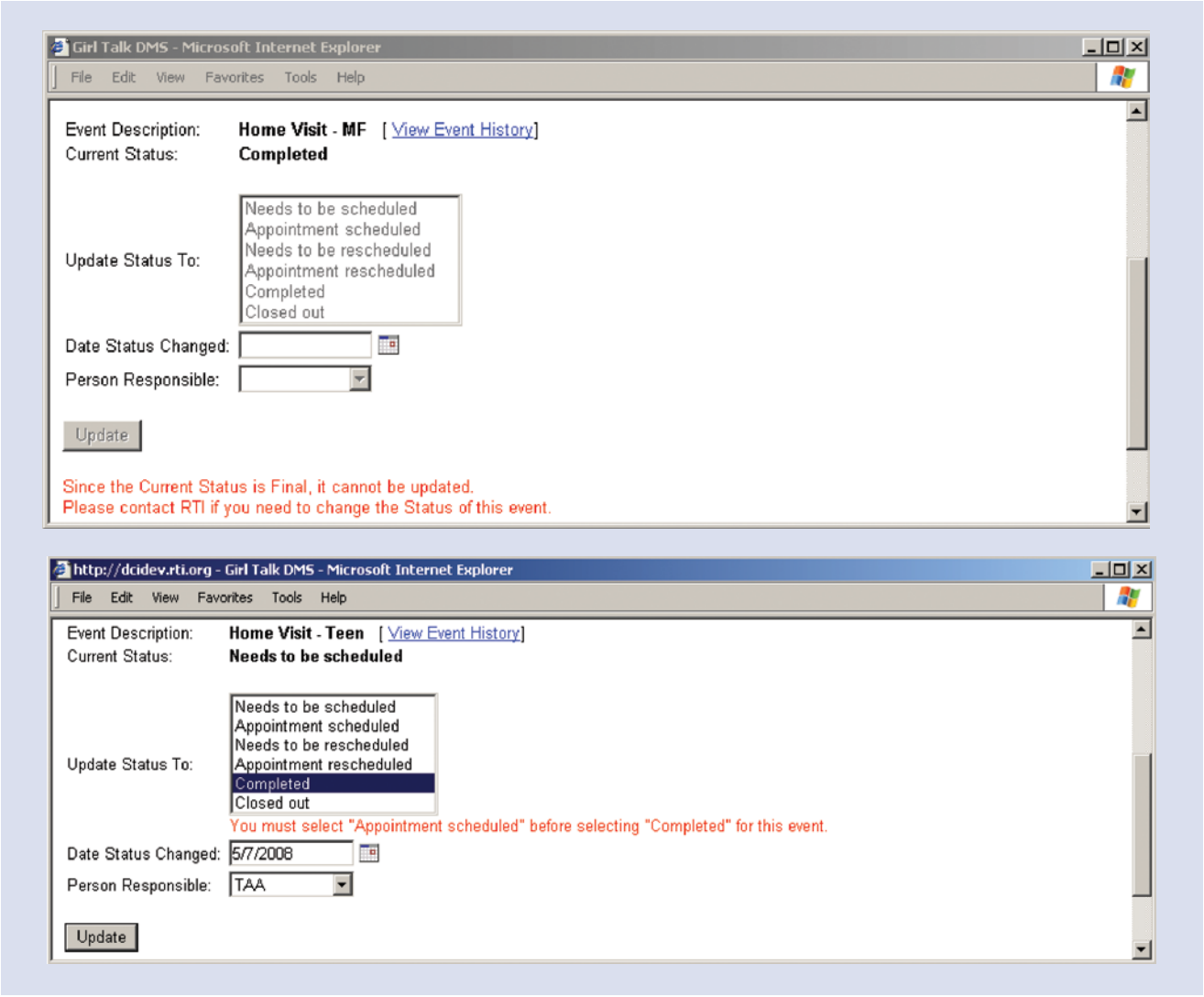
Events Tracking QA/QC Features

Because the events tracking system is critical to study success, we have added several types of QA/QC checks to ensure that staff complete all study activities required by the protocol. Figure 6 shows an image of the events tracking QA/QC features.

For QA/QC purposes, the DMS prevents a user from changing any event status that is considered a final event status (e.g., "consent granted," "incentive mailed," or "completed"). If a user selects a final status incorrectly, he or she contacts RTI's DMS administrator to reset it. This prevents study staff from inadvertently reversing a status that may have triggered additional events. When resetting a status, the DMS administrator removes any additional events or status changes that had been triggered.

We have added additional QA/QC features to prevent counselors from skipping a required status. Because the DMS triggers additional events when counselors schedule the home visit to conduct the baseline interview, "scheduled" is a required status

Figure 6: Events tracking QA/QC features



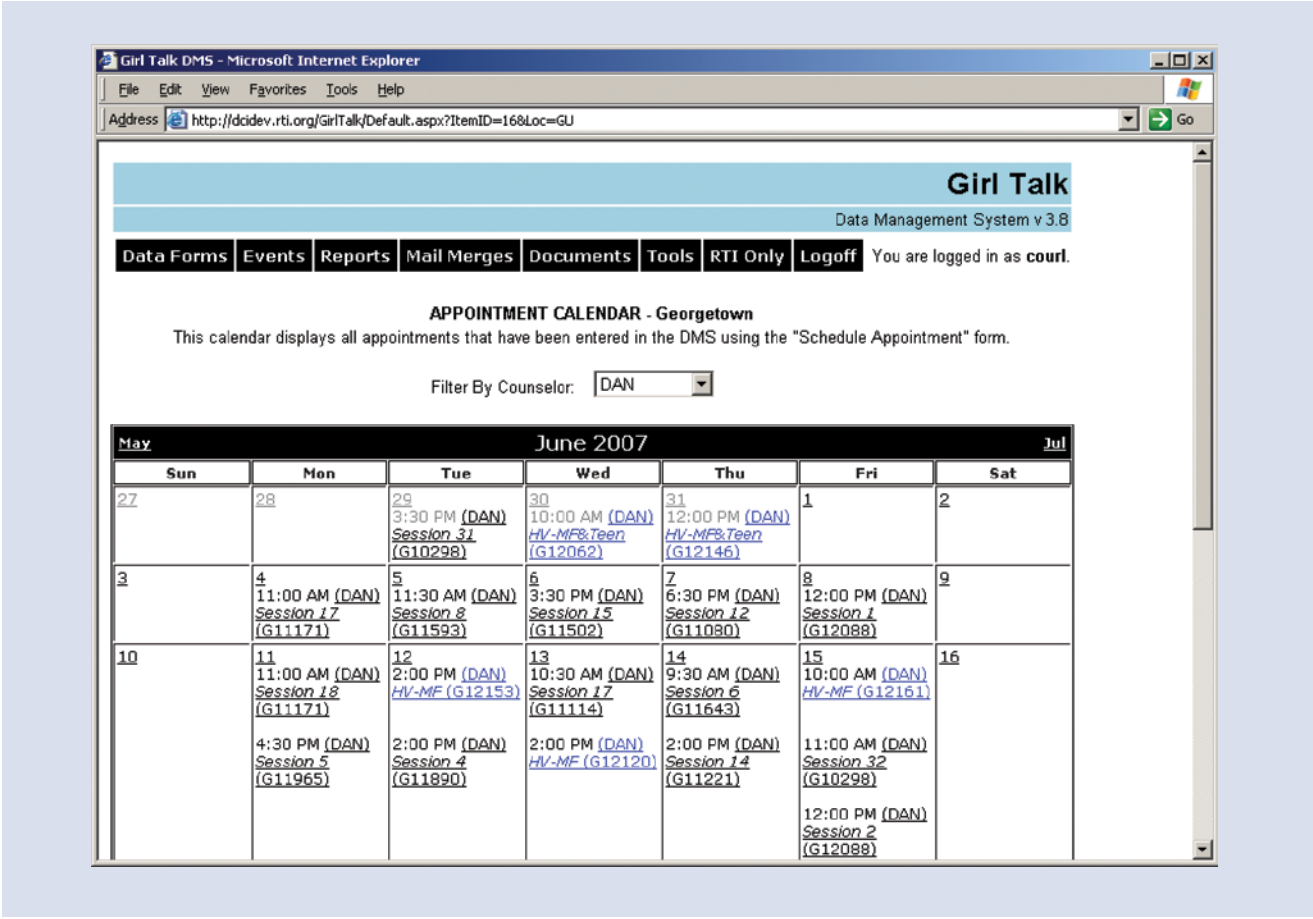
for the home visit. Thus, the DMS QA/QC checks prevent counselors from selecting "completed" before selecting "scheduled."

Appointment Calendar

Each location has its own appointment calendar. The one shown in Figure 7 is for GU. Staff can filter appointments to show those assigned to a particular staff member. When the user clicks the link for an appointment, a small window pops up with details about the appointment. The pop-up window includes a MapQuest link to the appointment location, so that users can view a map of the appointment's location.

The calendar also displays appointments for other protocols in the NIH-DC Initiative. For example, in addition to conducting evaluation interviews for GirlTalk, interviewers at CNMC conduct interviews for two other NIH-DC Initiative studies. Appointments from the other studies, which are displayed on the GirlTalk calendar as read-only, allow interviewers to identify scheduling conflicts easily. A calendar could be programmed in a PC-based system; however, the data from other sites would not be in real time.

Figure 7: Appointment calendar



Report Generation

Staff can generate reports in the DMS using real-time data from all sites. The DMS report "Pregnancy Test Reminders and Self-Check Calls," shown in Figure 8, is an example of how information from different components feeds into one report. Interviewers at CNMC call the teens about their pregnancy tests, but the counselors at GU are responsible for keeping the teen's primary telephone number up-to-date. When the counselors update the phone number, the DMS instantly reflects this change on this report. Thus, from this report, the CNMC interviewers have the latest contact information for the teen.

Most of the reports use a certain Event and Status to determine which records to display (for example, "Pregnancy Tests" that are "Needed"). In addition to customized reports, the DMS contains an interactive report known as the "Create Your Own Event Report," which allows users to generate a report dynamically for any event and status code to which they have access. This feature substantially reduces the need for programmer-developed custom reports.

Study Documents

Study documents, such as consent forms, showcards, and questionnaires, are stored centrally in the DMS to ensure that users always access the most recent version of a document. Figure 9 shows the Documents feature. With a PC-based system, if study staff update or add new documents, RTI would need to transmit them to each PC and the field staff would need to save them to their computers' hard drive.

Figure 8: Report generation example

Uploading Interview Data

The counselors and interviewers use the DMS to upload ACASI and CAPI data files to the server at the DCC, eliminating the need for a separate data transmission system that would be necessary with a PC-based DMS. After staff upload the data files, an automated process moves the interview records into the central study database. At this point users can view selected ACASI and CAPI data items on reports in the DMS.

The DMS also contains Data Uploads reports showing which interviews have been uploaded for each teen. Figure 10 shows an image of the Upload Files page.

Figure 9: Study documents

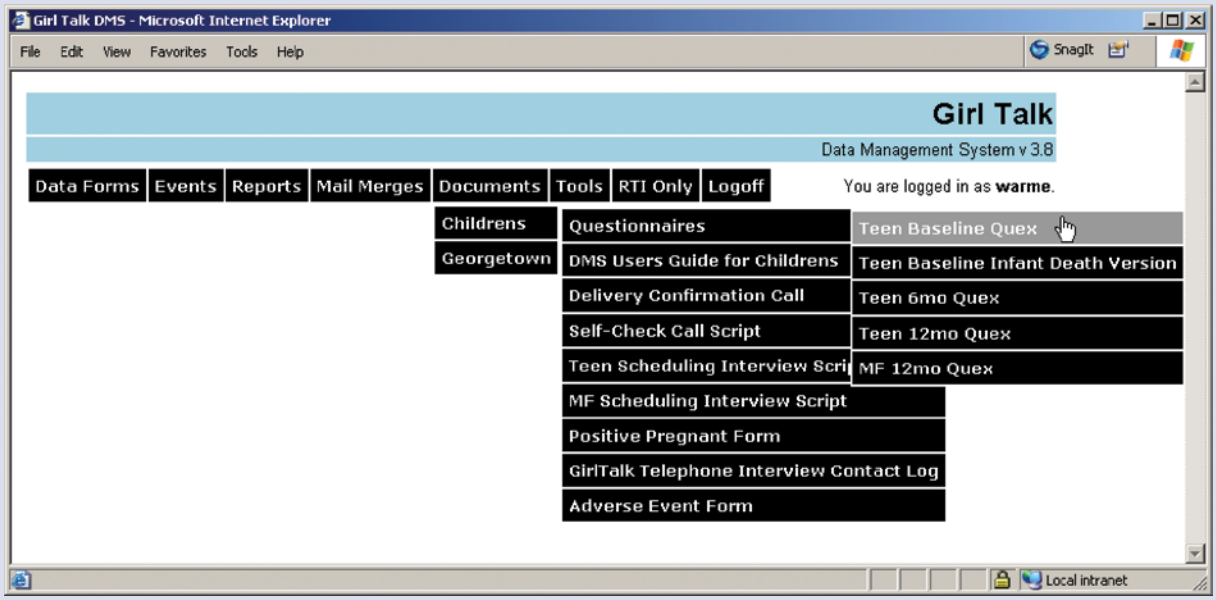
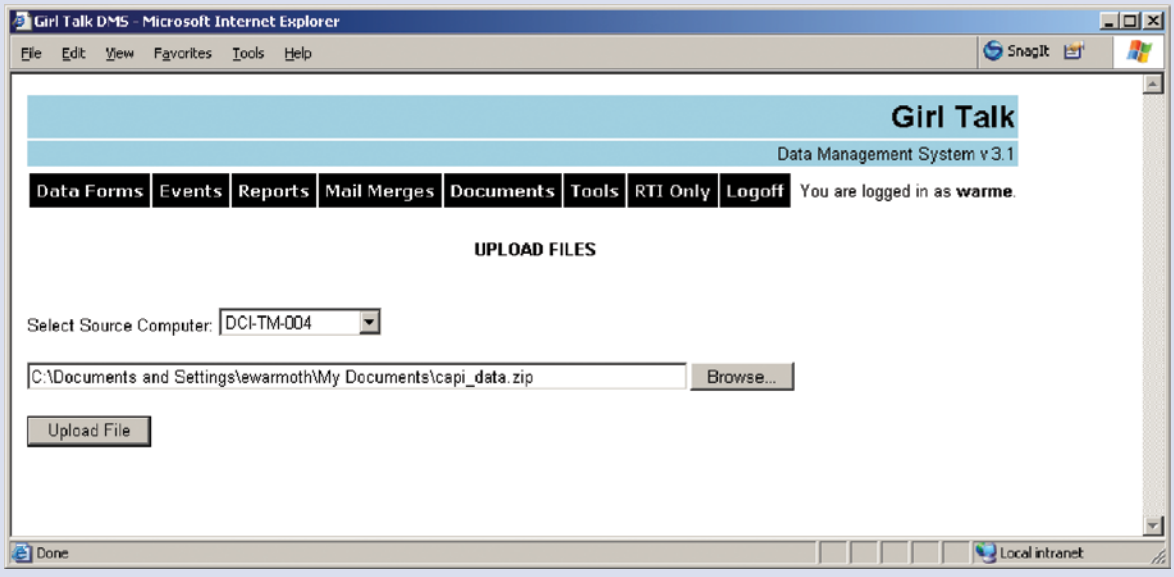


Figure 10: Interview data file upload



Mail Merges

Because staff frequently send incentive mailings to each participant during her enrollment, the DMS generates necessary cover letters with prepopulated name, address, and date. When the Events section indicates that a letter is needed, the DMS creates the letter, which is printed by counselors. Programmers could implement mail merges with a PC-based system, but the data may not be up-to-date.

Participant Notes

A large free-form text box is available for GirlTalk staff from all sites to add notes for any participant. The comments fields are shared between sites and serve as a place to incorporate and exchange key information for each participant. Programmers could implement this in a PC-based system, but the programmer would need to use a different design to avoid overwriting participant notes when transmitting and syncing data between different sites.

Discussion of Lessons Learned

In April 2005, development of the DMS began, with the first DMS staff training conducted in July 2005. RTI launched the DMS in three stages: Stage 1, released in September 2005, contained features for Recruitment. Stage 2, released in November 2005, contained features for Evaluation. Stage 3, released in February 2006, contained features for the Intervention.

We have released a total of 13 versions of the DMS since its inception. Each version incorporated new features requested by field staff during each stage and changes to accommodate procedure modifications requested by study investigators. As of September 1, 2008, 25 DMS users had logged more than 12,750 sessions to record study activities and data. Study staff had screened a total of 286 participants and randomized 223. The DMS had generated more than 7,260 events for study participants, with an additional 4,070 events scheduled as due in the future.

In the next sections we discuss the lessons learned from our experiences in developing and implementing the DMS.

Design Lessons Learned

Efficient data management that fulfills client and user expectations is necessary, especially for research studies that have budget limitations. Even with the most complex studies, maintaining a simple system is the best approach.

To reduce labor and programming time, study managers should complete as much of the study planning as possible before the DMS programming begins. Advance preparations will reduce labor costs and ensure that an organized framework for the DMS is created. All required data entry forms should be conceptualized before programming the DMS. This will enable study staff to create a more concise structure, which will reduce the quantity of forms needed. In addition, all study procedures to be monitored by the DMS events tracking system should be finalized in advance to keep the structure simple and reduce the need for reprogramming.

Any modification to the event tracking added at a later date may cause complications because of its impact on other events. One disadvantage of the events tracking system is that if researchers or data managers add a new event mid-study, then the update will not affect the participants who have already passed the trigger point. For example, the DMS triggers all follow-up interviews at the time of randomization. Thus, if the study adds new events after randomization has occurred, the programmer will need to trigger the new events manually.

Before any DMS of this sort is put into the field, we highly recommend a pilot test. Because of time constraints, we did not pilot-test the GirlTalk DMS. When the study experienced some procedural changes in the first few months of the study, these affected the DMS, which required additional reprogramming.

The sophisticated Web-based DMS contains enhanced features that allow for some automatic triggers, as discussed previously in Events Tracking QA/QC Features (p. 10). However, users tend to expect the DMS to conduct all processes automatically, which is not the case. Users must ensure that all data entered into the DMS are correct;

if data input into the DMS are incorrect, the DMS will also be incorrect. Because multiple users access data in the DMS, double-checking accuracy is especially critical.

Although the DMS eliminates much need for continual communication between sites, the GirlTalk study recognized the importance of holding weekly or biweekly multisite staff meetings. We established the weekly meetings to discuss common issues faced by all sites, in addition to addressing potential errors or questionable data in the DMS.

Implementation Lessons Learned

A phased approach to implementing a system works well when the development time frame is short, as it was for GirlTalk. The study team determined which features users initially needed and programmed those first. Then RTI updated the DMS incrementally as the study progressed and other features were needed.

A flexible DMS design should be balanced with features that are automated and controlled by the DMS. Although users often seek flexibility, maintaining data quality requires that the DMS have some rigidity. For example, because randomization is critical, it must be controlled by the DMS. Before allowing staff to randomize a participant, the DMS confirms that all required baseline data have been entered. In addition, the DMS prevents randomization from occurring more than once per teen.

The Web-based DMS facilitates entering clean data. As compared with the PC-based DMS used in DC HOPE, the Web-based DMS produced data that did not need to be cleaned as frequently. In addition, the Web-based design eliminated transmission difficulties so that the data were available in real time. The RTI team did not maintain data reports to enumerate these differences, but we observed them and recorded them anecdotally.

Another lesson learned was that if a software tool with the desired functionality already exists, study teams should consider using it. For the GirlTalk project, users requested complex appointment scheduling features. RTI recommended that staff use Microsoft Outlook for complex scheduling needs and use the DMS for more basic scheduling functions. Ultimately the DMS users decided to use the basic scheduling features programmed in the DMS.

Finally, considering all requests from users and taking advantage of their input to build a better system is vital. Users brought forth the idea to add a calendar to the DMS, which we originally thought to be a complex enhancement requiring resources that were limited. However, with brief investigation, we found the calendar feature to be easy and quick to implement.

Cost Considerations for a Web-Based System

The initial design and development of a Web-based DMS can be expensive. The following are some factors to consider in determining whether the study can justify the cost:

- Size of the study—The size of the study can refer to the number of cases, the number of study activities, and/or the amount of data being collected. The larger the study, the lower the cost per case or activity.
- Duration of the study—If the study runs over a long period of time, such as several years, then the investment per year will be reduced.
- Multiple sites—A Web-based DMS is especially suited for multisite studies because it enhances site-to-site communication by providing users from all sites with real-time data.
- Infrastructure—Technical support costs are associated with maintaining the Web server, database server, and Internet connections. Study managers should weigh these costs against the costs associated with maintaining a PC-based system in the field. Project leaders need to consider who will be providing the technical support and what it will cost in both the Web- and PC-based scenarios.
- Reusability—Once a Web-based system is developed for one study, reusing it (or adapting it) for other studies may be possible. The potential opportunity to recycle the DMS infrastructure across different studies may justify the initial investment in developing the system.

Conclusions and Outcomes

The Web-based DMS creates an efficient and effective management tool for complex, multisite studies. Customized reports allow managers to monitor study progress and conduct regular quality control reviews to ensure data quality. Providing staff with real-time access to all data reduces the volume of communication required between sites. However, regular communication between sites and among study staff continues to be a necessary aspect of study management, allowing study staff to confirm or question data inputs. We emphasize that the accuracy of the system depends on the accuracy of the data entered by the user.

Using a Web-based DMS alleviates many of the problems encountered with a PC-based system. For example, a Web-based DMS reduces transmission difficulties. Real-time access to data eliminates issues of sharing data between sites.

The Web environment allows for new features that are difficult or impossible to implement with standalone PC systems. For example, centralized events tracking, appointment scheduling, and document storage are not possible in the conventional PC-based DMS.

The Web-based DMS is a cost-effective investment for studies requiring complex management and a large sample size or lengthy longitudinal studies. Once the DMS has been developed, other programmers can reuse the DMS infrastructure and portions of the code to create new DMS applications for additional studies.

Future Directions

Future modifications for the Web-based DMS include building new tools to speed the development of DMS applications. For example, RTI would like to create a rapid development tool for the data entry forms. This tool would use a codebook containing information about the data items on a form and generate the data entry forms automatically.

A second enhancement is to manage the event specifications within the DMS instead of in Microsoft Excel. Currently RTI creates, manages, and maintains all event specifications in an Excel spreadsheet and loads them manually into the event tables in the study database. Programmers could add features allowing RTI staff to create events and status codes, associate certain status codes with each event, and specify which events trigger other events. This feature would directly modify the event tables in the study database, thereby eliminating the burden of maintaining the event specifications in two different formats and keeping them in sync.

RTI migrated the GirlTalk DMS and the other NIH-DC Initiative DMS applications to ASP.NET version 2.0 midway through the study. ASP.NET 2.0 offers many new features that developers had to design and build themselves in version 1.1. The Web-based DMS will evolve to take advantage of the latest capabilities. For example, the new Site Navigation piece of ASP.NET 2.0 allows a developer to define the navigational structure of the website. These data can then be used by a variety of Web controls to display site maps, breadcrumbs, tree views, or menus that highlight the site's navigation and show the user's location in the site.

Also, the new ASP.NET AJAX extensions allow the developer to build Web applications in ASP.NET 2.0 that can update data on the Web page without a complete reload of the page (a "round trip" to the server). The user may not even perceive that a page update is occurring and can continue interacting with the page while the update is happening in the background. This capability will enhance the richness and responsiveness of the DMS application.

References

Miniwatts Marketing Group. (2008, June). Internet world stats: usage and population statistics. Internet usage statistics for the Americas, according to Nielsen NetRatings Inc. (June 2008). Retrieved September 9, 2008, from http://internetworldstats.com/america.htm#us

Kochanek, K.D., Murphy, S.L., Anderson, R.N., and Scott C. (2004). *Deaths: Final Data for 2002.* National Vital Statistics Reports, Vol. 53, No. 5. Hyattsville, MD: National Center for Health Statistics.

Munson, M.L., and Sutton, P.D. (2006). *Births, Marriages, Divorces, and Deaths: Provisional Data for 2005.* National Vital Statistics Reports, Vol. 54, No. 20. Hyattsville, MD: National Center for Health Statistics.

Unutzer, J., Choi, Y., Cook, I., and Oishi, S. (2002). A web-based data management system to improve care for depression in a multicenter clinical trial. *Psychiatric Services, 53*(6): 671–673, 678.

Winget, M., Kincaid, H., Lin, P., Li, L., Kelly, S., and Thornquist, M. (2005). A web-based system for managing and coordinating multiple multisite studies. *Clinical Trials, 2*(1): 42–49.

Appendix A
Project Summaries

GirlTalk for Teen Moms Study

The major objective of the *GirlTalk for Teen Moms Study* (GirlTalk) is to test the effectiveness of a community based intervention for minority teen mothers with the primary purpose of increasing the interpregnancy interval to greater than 24 months. The secondary aim is to improve knowledge, attitudes, behaviors, relationships, and connectedness generally understood to be related to length of interpregnancy intervals.

Subjects are recruited from the population of pregnant teens residing in the Washington, DC, metropolitan area through a multitude of community sites, including schools, community centers, and health clinics. Teens are screened by project staff using the following criteria: African-American or Latina primiparous pregnant teens; English speaking; and age 15 to 18 inclusive or teens who are 19 years old if they have not graduated from high school.

Once eligible participants have been identified and guardian consent provided, they are randomized into control (usual care) and intervention groups. The intervention consists of a curriculum with both group-based and individual-focused components. Intervention subjects are provided with cell phones for maintaining contact. The project counselor provides individual interventions by cell phone over the 24-month study period.

Evaluations of the teen mothers are conducted in their third trimester of pregnancy (prenatal), within 2 to 3 weeks after delivery (postpartum), and at 6, 12, 18, and 24 months after delivery to assess the efficacy of the intervention. The mother/mother-figure completes a brief interview at baseline and at 12 and 24 months after delivery. An in-person home visit is scheduled to conduct both the prenatal teen interview and the mother/mother-figure baseline interview. All other interviews are conducted by an interviewer by phone. In addition, participants are evaluated for delay of a subsequent pregnancy. Pregnancy status is confirmed by urine testing at 6, 12, 18, and 24 months after delivery and by self-report at 3, 9, 15, and 21 months after delivery.

Given the number of teen pregnancies in Washington, DC, and accessibility to the major clinics where teens receive prenatal care, as well as multiple community sites, it is expected that a total of 340 teen mothers can be recruited over the 24-month period. The accrual rate is expected to be in the range of 14 to 15 teens per month. Recruited teen mothers will be followed through 24 months after delivery. Recruitment and screening in the study started on September 9, 2005. Data collection began on October 19, 2005.

Project DC HOPE

Interventions for Risk Factors in Pregnant Women in Washington, DC: An Integrated Approach, known as Project DC-HOPE (DC-Healthy Outcomes of Pregnancy Education), was one of the NIH-DC Initiative's Phase 2 studies. The study was implemented in the field in July 2001. Recruitment continued through October 2003; data collection, intervention, and follow-up activities concluded in April 2004.

The objective of Project DC-HOPE was to test the effectiveness of a community-based intervention for pregnant African-American and Latina women that aimed to reduce the prevalence and severity of two specific risk factors linked to adverse pregnancy outcomes—psychosocial risk and smoking—by providing health behavior counseling to pregnant African-American and Latina women in Washington, DC. A secondary goal of the integrated intervention program was to improve pregnancy outcomes as measured by infant birth weight and gestational age.

Pregnant women of African-American or Latina race/ethnicity who were Washington, DC, residents and at least 18 years old were eligible for this project. They were recruited at participating prenatal care clinics through 28 weeks' gestational age in their pregnancies. Pregnant women presenting at the clinic

sites were screened, using audio computer-assisted self-interviewing (A-CASI), to assess their eligibility for the study. Additional eligibility criteria included psychosocial risk (depression, partner abuse) and smoking (including exposure to environmental tobacco smoke). Eligible women who consented to participate in the study completed a baseline evaluation questionnaire.

Women assigned to the intervention group met with a trained pregnancy advisor at each prenatal visit and at two postpartum sessions to receive individualized counseling targeting their declared risk factors. They were evaluated periodically by the pregnancy advisors for all risk factors not previously declared and, if additional risks were identified, received intervention for these as well. Women assigned to the usual care group met with their primary care providers according to standard clinic practice.

Components of the integrated intervention targeting depression were based on the Cognitive-Behavioral theoretical model, and intervention content for partner abuse focused on developing a safety plan. The Stages of Change model provided the framework for the smoking cessation intervention. A brief educational component focusing on healthy reproductive practices was also provided to all women in the intervention group.

Appendix B

List of Data Management System Events

Event	Intervals	Status Codes
Screening Form		Completed—No LG
		Completed—Has LG
MF Identification by Teen		Needs to be completed
		Completed—No MF
		Completed—MF
		Closed out
MF Identification by LG		Needs to be completed
		Completed—No MF
		Completed—LG = MF
		Completed—LG <> MF
		Closed out
LG Written Consent — LG <> MF or LG Written Consent — No MF		Needed
		Mail merge produced
		Mailed
		Consent granted
		Consent refused
		Closed out
LG Written Consent — LG = MF		Needed
		Consent granted
		Consent refused
		Closed out
LG Written Consent Reminder Call		Needed
		Completed
		Closed out
Home Visit—Teen and Home Visit—MF		Needs to be scheduled
		Appointment scheduled
		Needs to be rescheduled
		Appointment rescheduled
		Completed
		Closed out
Teen Assent/ Consent		Needed
		Assent granted
		Assent refused
		Closed out
MF Informed Consent		Needed
		Consent granted
		Consent refused
		Closed out
Teen Photo and Release Statement		Needed
		Signature by LG obtained
		Signature by teen obtained
		Signature by LG+teen obtained
		Photo + signatures obtained
		Consent refused
		Closed out
Teen Prenatal Interview, MF Baseline Interview, and Teen ACASI		Needed
		Partially completed—needs to be rescheduled
		Partially completed—rescheduled
		Completed
		Refused
		Closed out
Incentive Forms—Teen and Incentive Forms—MF		Needed
		Completed
		Closed out
Teen Prenatal Incentive and MF Baseline Incentive		Needed
		Mail merge produced
		Mailed
		Closed out
Delivery Check-in Call		Needed
		Completed—baby not delivered
		Completed—baby delivered
		Closed out
Teen Postpartum Baseline Interview		Needs to be scheduled
		Appointment scheduled
		Partially completed—needs to be rescheduled
		Partially completed—rescheduled
		Completed
		Refused
		Closed out
Randomization		Needed
		Completed—Usual Care Group
		Completed—Intervention Group

Event	Intervals	Status Codes
Baseline Incentive —Usual Care and Baseline Incentive—Intervention		Needed Mail merge produced Mailed Closed out
Teen Evaluation	6, 12, 18, 24 months	Needs to be scheduled Appointment scheduled Partially completed—needs to be rescheduled Partially completed—rescheduled Past due event Completed Completed short version—repeat pregnancy Lost contact Refused Closed out
MF Evaluation	12 months, 24 months	Needs to be scheduled Appointment scheduled Partially completed—needs to be rescheduled Partially completed—rescheduled Past due event Completed Completed short version—repeat pregnancy Lost contact Refused Not needed—No MF Closed out
MF Evaluation Incentive Teen Evaluation Incentive	12, 24 months 6, 12, 18, 24 months	Needed Mail merge produced Mailed Closed out
Pregnancy Test	6, 12, 18, 24 months	Needed Past due event Completed—not pregnant Self report only—not pregnant Doctor confirmation needed—pos pregnant Completed—pos pregnant EPT not needed—pos pregnant confirmed Closed out
Pregnancy Test Reminder	6, 12, 18, 24 months	Needed Completed Closed out
Pregnancy Self—Check Call	3, 9, 15, 21 months	Needed Past due event Completed—not pregnant Completed—pregnant or DK Completed—confirmed pos pregnate Closed out
Pregnancy Test Confirmation	3, 9, 15, 21 months	Needed Completed—not pregnant Doctor confirmation needed—pos pregnant Completed—pos pregnant Closed out
Showcards—Teen/MF	6, 12, 18, 24 months	Needed Mailed Closed out
Cell Phone Delivery		Needed Completed Refused Closed out
Intervention Session	45 sessions	Needs to be scheduled Appointment scheduled Completed Closed out
Intervention Dinner Group	8 meetings	Needs to be scheduled Teen notified of meeting Missed—needs to be rescheduled Teen notified of makeup meeting Completed Closed out

ACASI = audio computer-assisted self-interviewing
DK = don't know
EPT = early pregnancy test
LG = legal guardian
MF = mother figure
pos = positive

Acknowledgments

This study was supported by grants 3U18HD030445, 3U18HD030447, 5U18HD31206, 3U18HD031919, and 5U18HD036104 from the Eunice Shriver Kennedy National Institute of Child Health and Human Development. The authors would like to acknowledge the support and assistance of Jutta Thornberry, BA, Nabil El-Khorazaty, PhD, David Forvendel, BS, and Margo Brinkley MS, from RTI International, as well as the GirlTalk staff at Georgetown University and Children's National Medical Center.

www.rti.org/rtipress **RTI Press publication MR-0007-0812**