

6.1 Software Requirements for Daily Operations

6.1a Software Provided by CDC. Several software applications are required for conducting daily PRAMS operations. Of these, CDC provides the following four applications at no cost to state PRAMS projects.

i. PRAMS Integrated Data Collection System (PIDS) is the centerpiece of the applications used for daily operations. It is used to schedule and track data collection activities, and record data on mail and telephone results. It is also a reporting program that can be used to list or summarize current operational data in order to facilitate daily operations. It is also used for telephone follow-up. It manages call attempts for telephone interviews. It will keep track of scheduled appointments, manage multiple phone numbers, prioritize cases to call, assign current status codes for each mother, and generate phone operations files. It records phone survey responses and extraneous comments provided by mothers. PIDS integrates the following software systems to conduct operations.

a. SugarCRM Word Plug-in is used to produce personalized mail correspondence using the mail merge capabilities of a word processor.

b. IBM SPSS Data Collection Interviewer is used to record mothers' responses from mail questionnaires. It is also used to record mothers' responses from "late" mail questionnaires that have been returned after a batch is closed, and to record the operations information for these "late" mail questionnaires. Also, it is used to record mothers' miscellaneous written comments that fall outside of the standard questions and answers of the mail questionnaire.

ii. Secure Access Management System (SAMS) provides identity proofing and authentication services. It also provides secure file transfer facilities for sending pass through files and receiving certified data files. These files are secured with an address and password.

6.1b Software Provided by the State. States are responsible for providing software applications that are used for conducting daily operations and monitoring phone interviews:

i. Microsoft Office 2003 or higher (Word is required). SugarCRM Word Plug-in works in conjunction with a word processor to produce personalized letters that are a part of each mailing.

6.1c Web-based Analytic Tools Provided by CDC. CDC provides the following tools at no cost to state PRAMS projects.

- i. **PRAMStat** is a public use web-based query system which includes most core and selected standard indicators from the PRAMS survey. A minimum 70% response rate is required for state data to be included in CPONDER from 2000 to 2006, and a minimum response rate of 65% is required from 2007 forward. CPONDER allows display of indicators across states and years. The data runs are based on a system of pre-programmed analyses generated using SUDAAN. In order to protect the confidentiality of participants in PRAMS, only cells where the marginal totals have at least a sample size of 30 are included in the outputs, and only 1 break-out variable at a time may be specified.

6.2 Computer Platform Requirements

In general, PIDS requires a standard entry level personal computer that is capable of running the latest version of the Windows Professional Operating system (currently Windows XP), Internet Explorer 8.0 or higher and MS Office 2000 or higher. As time passes, the standard configuration for an entry-level desktop personal computer will change with the operating system requirements.

6.2a Basic Requirements. CDC has developed the following minimum recommended computing platform for computers using PRAMS software:

- A personal computer with at least Pentium®-class processor operating at 1 GHz (2 GHz or faster recommended)
- At least 512 MB of RAM (1 GB or more recommended)
- At least 200 MB of available hard disk space
- Microsoft Internet Explorer 8 or above
- Minimum 16-bit color quality setting
- A graphics adapter with 1024 x 768 resolution or higher
- Microsoft® Windows® XP Professional (x86 32-bit edition) with Service Pack 3, Microsoft Windows Vista Business or Enterprise Edition with Service Pack 2 (x86 32-bit or x64 64-bit edition), or Microsoft Windows 7 Business or Enterprise Edition with Service Pack 1 (x86 32-bit or x64 64-bit edition)
- Microsoft Word 2003 or 2007

6.2b Additional Requirements. CDC recommends the following additional equipment at the state:

- A laser printer that prints on envelopes, letterhead, label sheets, and traditional letter- and legal-size paper and that has a correctly functioning printer driver for the operating system being used. [In general, laser printers that are less than three years old and that are made by one of the mainstream printer manufacturers (such as Hewlett-Packard, Lexmark, Okidata) will meet these requirements.]
- High speed Internet connection for all computers
- Phone headsets for all interviewers

6.3 Virus Protection

All computers used for PIDS should have anti-virus software installed. State PRAMS staff should use anti-virus software to scan disks, e-mail attachments, and files downloaded from the Internet. Also, anti-virus software needs to be kept up-to-date by acquiring new virus definition files when the software manufacturer makes these available. Most state health departments have requirements for the use of anti-virus software. If not, CDC can make specific usage recommendations.

Protocol Development Task

Specify the brand name and version of the computer virus protection software to be used to protect your computer and data diskettes from viruses.

6.4 Data Security and Personal Identifiers

The software applications used for daily PRAMS operations contain substantial personally identified information. For this reason, the following security precautions exist:

- A screen saver login is required.
- PIDS documents user login activity.

Some computer-generated reports and other electronic and paper documents contain personally identified information. Therefore, PRAMS states must implement physical security to protect these files and documents. Physical security should include the following precautions:

- Files stored on a network should be in directories accessible only by PRAMS staff.
- Computers should have password-protected screen savers so that unauthorized people cannot use a computer that has access to PRAMS files.

- Completed questionnaires should be stored in locked cabinets.
- Completed questionnaires and other printed documents that contain personally identified information should be destroyed (shredded or burned) and not simply thrown away or recycled when they are no longer needed. See **Section 10.4** for your state's policy on the archival and destruction of records.
- When a computer used for PRAMS is taken out of service, any hard drives that may have once contained PRAMS data should be reformatted before being used for another purpose.
- Floppy disks and other removable storage media that are no longer needed for PRAMS should be destroyed and not used for another purpose.

PIDS also houses substantial personally identified information and extensive security precautions have been put in place to protect the confidentiality of this information. (see **Appendix V, PIDS Data Security Plan**). There are additional security measures that state staff can implement to further protect the data. These measures include the following:

- Staff logged into the PIDS system should log out before leaving their desk, even if just for a short break.
- PIDS administrators should promptly delete user accounts for state or contract staff that leave the project.
- Staff with access to PIDS should memorize their login information and never share it with anyone.

Protocol Development Task

Describe in detail all security precautions developed to protect access to PRAMS computer and paper files. If using a LAN, provide details on LAN security precautions.

Starting in 1999 and annually every since, CDC's Institutional Review Board (IRB) reviews the PRAMS protocol and other documents in order for PRAMS to receive IRB approval. As a result of this process, the CDC PRAMS team has implemented procedures for taking greater care when handling personal identifiers (e.g., names, birth certificate numbers, social security numbers), either in final data, during weighting, or when performing software support. The following guidelines have been implemented:

- Birth certificate numbers are no longer used as the key index for any application's database or program. This reduces the number of reports and files that contain birth certificate numbers along with other identifiers.

- CDC and the CDC contractor (which provides support for PIDS) will make every effort to resolve software support problems without asking states to send database files using electronic mail to CDC or the CDC contractor.
- Despite the best efforts of CDC and the CDC contractor, some software support situations will require access to state data. CDC and the CDC contractor have identified only a few individuals who may need to access personal identifiers for the reasons stated above. Access to files with personal identifiers is limited to those individuals.

6.5 Making Batch Data Available to CDC

PIDS maintains and produces the data for PRAMS surveillance and states can make this data available to the CDC each month. States will have the capability to expire completed batches at which point CDC may consider the batch ready to process. The types of data are discussed below, broken out into two groups; 1) Files that are imported into PIDS and 2) Data that is released to CDC:

These files are part of the PIDS import process:

- 6.5a Sampling Frame Birth Certificate Variable File (*stFR#.DAT*).** This file is created by <STATE> and will contain one record for every mother in the sampling frame (all records eligible for inclusion in the sample). The Sampling Frame file is used to evaluate the sample selection and to assess any bias that may have been introduced in drawing the sample. It does not become part of the PRAMS analysis data set. The variables and layout of the Sampling Frame file are identical to the variables and layout of the Sample Birth Certificate file to facilitate the assessment of selection bias. For an explanation of the creation of the Sampling Frame file, see **Section 4.4a**.
- 6.5b Sample Birth Certificate Variable File (*stBC#.DAT*).** This file is also created by <STATE> and contains information from the birth certificate that will be included in the analysis file. There is one record for each mother in the PRAMS sample (all records selected for PRAMS). These data will be used to assess response bias, verify demographic and other data, and analyze the relationships between maternal behaviors and birth weight and gestational age. Core birth certificate data items have been identified and incorporated into the Sample Birth Certificate file. <STATE> can select up to 25 additional columns of data to be included in this file, if desired.
- 6.5c Sample Contact Information File (*CONTACT.DAT*).** This file is created by vital records each month and is used to transfer contact information from birth certificate records into PIDS. It is important to note that this file differs from the monthly Sample Birth Certificate file in that it includes names, addresses, and several other variables that are required for tracking and contacting sampled mothers. Because of the personally identified information in this file, it is not

maintained as part of the monthly batch data, is not made available to CDC, and does not become part of the analysis data set. Because this file is imported into PIDS, it is important that this file be thoroughly tested by CDC. **See Appendix P** for guidelines on the monthly creation of the CONTACT.DAT file.

Protocol Development Task

Specify procedures for obtaining or downloading the Sampling Frame file, the Sample Birth Certificate file, and the CONTACT.DAT file from vital records to the PRAMS-dedicated computer each month.

Data Released to CDC with each Batch: PIDS has a batch expire process that allows an authorized state user to make the following data available to CDC (see the **PIDS User Guide** for instruction).

- 6.5d Operational Data.** This data contain information that summarizes the results of mail and telephone contact attempts. This information is important for monitoring and evaluating the ongoing surveillance methodology and the resulting response rates. The three types of data (Operations, Mail Details, and Telephone Details) are produced by the PIDS application that CDC provides. The Operations Output data becomes part of the analysis data set. The Mail Details and Telephone Details data can be used to evaluate the efficiency and quality of mail and telephone operations.
- 6.5e Questionnaire Data.** This data contain responses recorded from completed mail, web, and telephone questionnaires. This information provides the basis for the analysis data set. Mail questionnaire responses are keyed using the PIDS data entry that CDC provides; the program performs range and edit checks as data are entered. The program produces the Mail Questionnaire data that becomes part of the analysis data set. Web survey data is keyed directly into PIDS by the respondent. Telephone questionnaire responses are keyed into the PIDS system, which also performs range and edit checks as the data are entered. The Web and Telephone Questionnaire data are produced by the PIDS system and are made available to CDC along with the mail questionnaire data. The questionnaire data includes core items, common to all PRAMS states, as well as state-specific items that states have chosen to add to their questionnaires.
- 6.5f Comment Data.** This data contains any additional comments that were written on the questionnaire or that were expressed in a telephone interview. Write-in comments are often invaluable in clarifying or interpreting the response to a particular question. The Comment file can be linked to the questionnaire files so that the questionnaire responses and write-in comments can be viewed together. Using the Comment program that CDC provides, mail comments are entered verbatim, with one exception. No names, addresses, telephone numbers, or e-mail addresses are recorded for either the woman or her provider(s). The PIDS

system allows for the entry of telephone comments during the telephone interview and web survey comments during the web session.

The operational, questionnaire, and comment data is maintained in PIDS.. <STATE> is responsible for producing the birth certificate variable files (stFR#.DAT, stBC#.DAT) and the CONTACT.DAT file from state birth certificate records. **Appendix Q** provides uniform file layouts for all files except CONTACT.DAT; the uniform file layout for CONTACT.DAT is located in **Appendix P**.

Protocol Development Task

Your state is responsible for developing variable definitions for the state-specific portions of the **sample birth certificate file**. CDC will assign variable names. Describe the state-specific variables selected for the sample birth certificate file in **Appendix Q (Uniform File Layouts)**. See **Chapter 4 (Sampling)** for more information about the layout of the state-specific portions of the Sample Birth Certificate file.

Your state is responsible for selecting standard and/or state-developed questions to add to your questionnaire. CDC will develop variable definitions, assign variable names, and provide a codebook for the **questionnaire files**. Insert the questionnaire codebook provided by CDC in **Appendix Q (Uniform File Layouts)**.

6.6 Quality Control

- 6.6a Data Cleaning and Editing.** Data editing is the process of checking the batch data files for data entry errors and inconsistencies before submitting the files to CDC. Data editing should be done after questionnaire and comment data entry are completed and the batch is closed (completed) in PIDS. CDC has developed guidelines for the general cleaning and editing of all data files (see the **PIDS Implementation Manual** for instruction), and PIDS will produce a Compare Report to assist in this process.
- 6.6b Data Entry Verification.** The mail questionnaire data entry program provides automatic range checks. However, it cannot detect erroneous data that are within the range defined for those data items. The mail data entry module of PIDS has a procedure that allows double-entry of selected questionnaires to check for keying errors. It is necessary that 10% of all mail questionnaires be double-entry verified. These procedures are especially useful in evaluating new data entry personnel and for identifying “problem” questions that are particularly prone to keying errors. The Verification Workbench provides the verification results and allows users to view any discrepancies found during the verification process and to make corrections as needed. See the **PIDS User Guide** for guidelines for verification of mail questionnaire data entry and reviewing verification discrepancy statistics. The nature of the phone data entry module prohibits double-entry of telephone survey responses. However, phone staff may review

answers provided during phone interviews by using the review mode. In review mode, the form is populated with the answers that were originally entered during the interview and staff can make any changes as needed will automatically in review mode. It is not possible to perform verification of web survey responses. However, the web survey provides automatic range checks and checks for answers left blank. Skip patterns are automatically implemented based on the response to the filter question.

6.6c Telephone Interviewer Monitoring. CDC has developed procedures for monitoring the consistency and quality of telephone interviewing. A proportion (at least 10%) of calls made by each telephone interviewer should be monitored for each batch. The phone module of PIDS allows state staff to monitor phone interviews and observe real-time phone interview data entry. This feature does not provide any audio monitoring capability. This feature must be set up by the state. See **Section 5.8, Appendix M (Telephone Interviewer Monitoring Procedures) and PIDS User Guide**, for a more detailed discussion about interviewer monitoring.

6.6d Batch Checking Procedures at CDC. When a state completes data cleaning and editing and data entry verification and monitoring for a batch, it releases the batch to CDC. Data made available to CDC will not include personal identifiers such as names and addresses but will contain the identification number for each record. (The CONTACT.DAT file that is created each month and imported into PIDS is **not** available to CDC) Batch data is released electronically to CDC monthly (see the **PIDS User Guide** for instructions on releasing data). Once the data is released to CDC, the data are processed and consistency checks are performed.

6.7 Batch Reports

The PIDS report module allows states to create and save custom reports or to run

several pre-configured reports saved in the system. There are five pre-configured reports that are used by the Project Coordinator for evaluating and processing completed batches. These reports are also available to the CDC. The five reports are as follows:

- ActivityTimingReport
- AgeReport
- MailContactReport
- TelephoneContactReport
- ParticipationSummaryReport

Finally, the Summary Monitoring Report (see **Appendix M [Telephone Interviewer Monitoring Procedures]** for details) summarizes the monitoring efforts for the batch and is made available to CDC once a batch has been released. The report form itself can be produced in PIDS and must be completed by the Project Coordinator or other appropriate staff person, such as the supervisor of a survey research laboratory, before releasing the batch to CDC.

Protocol Development Task

In **Sections 6.5a-6.5f**, in the table in **Section 6.6d**, and in **Section 6.7** above, replace the “st” segment of the file and report names with your own two-letter state abbreviation. For example, Alabama’s Sampling Frame Birth Certificate Variable file name would be ALFR#.DAT.

6.8 Creating the PRAMS Analysis File

Annually, CDC will create for <STATE> a master analysis data set from which all PRAMS analyses will be conducted, provided the criteria below are satisfied. This analysis data set contains birth certificate, operations, and questionnaire data, and includes weights, other variables required by SUDAAN, and additional computed variables. For specific analyses, <STATE> will create smaller subsets of this master file, selecting the particular variables of interest. For CDC to create the PRAMS analysis file, the following criteria must be met:

- 6.8a Successful File Linkage.** To permit linkage of the birth certificate, questionnaire, and operations files into the analysis data set, <STATE> needs a common identification number for each mother on each file. When the CONTACT.DAT file is imported into PIDS, PIDS assigns a unique identifier called MomID. MomID contains information about the year of birth, batch number, and state, but does not contain any personal identifiers. MomID is used as a key index variable for PIDS.
- 6.8b Proper Data Management Procedures.** <STATE> must have carried out all data management procedures properly. These procedures include proper usage of software, monthly cleaning and editing of data, questionnaire verification, monthly batch report generation, and releasing the batch to CDC at designated intervals.
- 6.8c Adequate Response Rates.** Prior to 2007 <STATE> must have achieved a weighted response rate of at least 70% for meaningful analyses to be undertaken. From 2007 forward, the response rate must be at least 65%. The weighted response rate indicates the proportion of women sampled who completed a survey, adjusted for sample design. Adequate stratum-specific response rates (at least 65% for 2007 forward) should be achieved to perform stratum-specific analyses. All states that follow sound operational procedures are

provided with a weighted dataset. However, only those with weighted response rates of 65% or higher should be used for analyses that will be presented outside the health department. We do, however, encourage all states to use their data internally for program development, evaluation, and collaboration.

6.8d Evaluation of Sampling Procedures. CDC must have evaluated <STATE>'s sampling procedures.

6.8e Sending of Final Birth File. <STATE> will send CDC the final birth file in requested format (currently the 1999 or 2003 NCHS format with some additional PRAMS and state-specific variables appended) for the calendar year. Each state is sent a data layout by the CDC that is fairly easy to implement (since each state already generates a similar file for the NCHS). This birth file is necessary to create the weighted data set that is used for data analysis. This file is compared with the PRAMS Sampling Frame files to identify records that were omitted from the frame that were in fact eligible for PRAMS. The state should reserve batch number "999" for the final birth file. The birth file should be submitted to CDC using SAMS. Birth files, once uploaded, are transmitted directly to the primary weighting statistician.

When the preceding criteria have been met, CDC will compute analysis weights. The analysis weight for each observation can be divided into three components, each accounting for a different factor. The adjustment components are:

- Sample Design
- Nonresponse
- Omissions from the Sampling Frame (i.e., Noncoverage of the Sampling Frame)

CDC will provide <STATE> with a written summary of the analysis weight computation each year. For more information concerning the computation and use of analysis weights, **see Appendix B.**

States generally have the analysis data set for a given calendar year 6-9 months beyond the birth year. This time frame is dependent upon the receipt of cleaned and edited data files and the final birth file from the state. CDC will provide this weighted data set for states to use approximately 1-2 months after receiving the final birth file. The birth file (and all PRAMS operational files and questionnaire data) must be received by CDC no later than December 1 of the year following the birth year to be included in the CDC multi-state surveillance report for that year. This birth file is accessible only by the CDC PRAMS weighting statistician and computer network personnel responsible for assigning data access-privileges. These parties have obtained CDC security clearance and have been trained in the handling of confidential and sensitive data.