Standard Operating Procedure for Database Operations

Environmental Health and Safety Division
RTI International*
Research Triangle Park, North Carolina

Prepared by: _Elezuz_   Date: _July 16, 2008_

Reviewed by: _Jame B. Hang_   Date: _7/16/08_

Approved by: _RKM Jayanty_   Date: _7-16-08_

---

* RTI International is a trade name of Research Triangle Institute.

RTI INTERNATIONAL

Contents

Standard Operating Procedure
for Database Operations

1.0 Introduction

1.1 Scope

This operating procedure covers database operation activities performed by program data processing staff. Data entry activities, such as Sample Handling Analysis Laboratory (SHAL) sample processing, are included in the SHAL standard operating procedures (SOPs).

1.2 Requirements

This procedure assumes a familiarity with general database concepts and the use of Microsoft (MS) Access and MS SQL Server programming tools, such as the Query Analyzer and Enterprise Manager. General MS Windows Server management skills are also assumed for supervisory personnel.

1.3 Hardware/Software Environment

Internal Server—RTI maintains an internal database server for use with the PM$_{2.5}$ Chemical Speciation Program. This server runs MS SQL Server version 2000 on the MS Windows Server 2003 operating system. Only internal RTI personnel are allowed access to this internal server (individual accounts are set up as described below).

External Server—An external server (i.e., on that is accessible from outside RTI's private network) is used to store monthly reports for review by the U.S. Environmental Protection Agency (EPA) and site data reviewers.

> **Note:** The names of specific forms, queries, reports, and programs to be run are italicized throughout this SOP.

2.0 Create Data Contact Account on the External Server

2.1 Summary of Task

This procedure describes the steps necessary to create a data contact account on the external server (geos1.rti.org). This procedure requires that the user have administrative rights on the external server.

2.2 Procedure

2.2.1 Use User Manager and select domain of geos1.rti.org.

2.2.2 Add user account and set password. Set password to not expire, not change on initial login, and not be changed by user.

2.2.3 Add the user to the PM$_{2.5}$ group.

2.2.4 Create a directory for user's data under the correct EPA directory.

2.2.5 Set security for the directory. Grant read/list access to the user, the Delivery Order Project Officer (DOPO), and the overall Project Officer (EPA01).

2.2.6 Send the data contact's account, directory, and password to the appropriate DOPO for transfer to data contact.

3.0 Request Domain Account for a New Non-RTI Temporary Employee

3.1 Summary of Task

Domain accounts can only be created by Information Technology Services (ITS) domain administrators. This procedure describes how to make a request to have a domain account created for a new temporary service employee (e.g., SHAL temporaries). Note that all RTI employees have a domain account created as part of the hiring procedure; therefore, this procedure is only required for non-RTI temporary workers.

3.2 Procedure

3.2.1 The Laboratory Supervisor goes to ITS Web site and completes an *Account Request Form*. Request ONLY NT Domain Account. Be certain to mark the employee as temporary on the form. The Laboratory Supervisor adds a notation to ITS to add the employee to appropriate groups (typically RCC_NT/PMSHALUsers).

3.2.2 ITS returns (rejects) the Account Request Form to Laboratory Supervisor, requiring the Center Director's approval.

3.2.3 The Laboratory Supervisor forwards returned e-mail to the Center Director for approval.

3.2.4 The Center Director approves the new domain account and forwards this approval to ITS.

3.2.5 ITS creates a new domain account and notifies the Laboratory Supervisor by e-mail.

3.2.6 The Laboratory Supervisor forwards account information to the Center Information Management Systems (IMS) Supervisor and the Database Supervisor.

4.0   Request Deletion of Domain Account for Terminated Non-RTI Temporary Employee

4.1   Summary of Task

Domain accounts may only be deleted by ITS domain administrators. This procedure describes how to make a request for deleting the domain account of a non-RTI temporary worker after his or her termination. Note that domain accounts for RTI workers are automatically deleted as part of their termination process; therefore, this procedure is only needed for non-RTI workers.

4.2   Procedure

4.2.1   The Laboratory Supervisor notifies the ITS Department, the EISD IMS Supervisor, and the Database Supervisor about the appropriate domain account to be deleted. Because all file and MS SQL Server access is through this account, this effectively removes their file and MS SQL Server access.

5.0   Add Employee to the SHAL Database Users Group

5.1   Summary of Task

This is task is performed for employees who need database access. This procedure requires administrative and MS SQL administrative rights on RTI's internal server. Only people who have domain accounts may be added to the database users group. New non-RTI temporary workers must have their domain account assigned (see Section 3.0) before they can be added to the users database group.

5.2   Procedure

5.2.1   Provide the ITS Department with the name and domain account of the person to be granted access to the SHAL database and the NT domain group(s) that he or she should be placed into (typically RCC_NT\PMShalUsers).

6.0   Process Delivery Order and Schedule Associated Sampling and Analysis Events

6.1   Summary of Task

This procedure describes the operations necessary to process an incoming delivery order and to schedule the associated analytical and sampling requests.

6.2   Procedure

6.2.1   Get delivery order information from the EPA DOPO.

6.2.2   Determine the information needed for delivery order processing from information provided by the EPA DOPO.

6.2.3   Enter information for delivery order into the database.

6.2.4   Run *delivery order form* (as a report) from the database.

6.2.5   Prepare the file folder for delivery order paperwork. Place the delivery order form into folder.

6.2.6   Identify each site on the delivery order and determine if it appears on RTI's list of sites with sampler and Air Quality System (AQS) information.

6.2.7   Determine if the site is listed on RTI's list of sites with sampler and AQS information.

6.2.8   Enter sampler and AQS information for the new site into the database, if necessary.

6.2.9   Determine the sampler type and analysis list needed for each site listed on the delivery order. (This assumes that the same type is required throughout the delivery order).

6.2.10   Use the lookup list to determine the sampling configuration needed for the selected sampling type and analysis list.

6.2.11   Determine the beginning and end dates for each site. Use the measurement request generation program to create measurement requests for each site and date combination.

6.2.12   Print the sampling request forms for the location and file them in processing folder(s).

6.2.13   Review the sampling forms to verify that scheduling is correct.

6.2.14   After all samples have been scheduled, set the delivery order status to requests scheduled.

6.2.15   Select delivery order for proofing.

6.2.16   Compare the summary report to delivery order and make changes, if necessary.

6.2.17   Update the SHAL schedule calendar to reflect additional workload.

7.0   Receive Data from Laboratory

7.1   Summary of Task

This procedure describes the receipt of data (in spreadsheets) from the analysis laboratories for direct import into the database.

7.2   Procedure

7.2.1   Receive the spreadsheet that contains analytical results from the laboratory.

7.2.2   Move spreadsheet data onto server into the appropriate laboratory file.

7.2.3   Review laboratory files to ensure that they are in the correct format for import. Make corrections to format as necessary for automatic import into database.

7.2.4   From the database, use the *import analytical data* form to automatically import analytical results into the database.

7.2.5   If errors occur during import, do not commit the transaction and identify and correct any problems with analytical data before importing the data.

7.2.6   Move imported laboratory results files into the added to database folder within each laboratory folder on the server.

8.0   Prepare Monthly Analytical Data Report

8.1   Summary of Task

This procedure describes the preparation of the analytical data report, which is sent to the EPA DOPO each month.

8.2   Procedure

8.2.1   Perform preliminary duplicate data check by running the *DignoseDuplicateRows* program.

8.2.2   Correct any duplicated data, as necessary.

8.2.3   Make a copy of the main database for use in report checking by running the *Transfer to QC Draft* program using the MS SQL Server Data Transformation Services.

8.2.4   Select the last sample for delivery date by setting the correct value for the last sampling date and by editing the date in the *ForceApproveTestBatch.sql* program.

8.2.5 Batch approve samples by running the SQL script by running the *ForceApproveTestBatch.sql* program.

8.2.6 Fix problems with reported uncompleted samples (in main database). Rerun everything to this point if any unaccounted sampling events remain.

The New_Export_Report_Data procedure performs the following calculations:

1. Ambient concentration = analyte mass/sampler volume (for appropriate sampler channel), where sampler volume is from the Field Custody Chain of Sampling form. If no sampler volume was supplied, but an average flow and elapsed time were supplied, then sampler volume = average flow * elapsed time.

2. Sample concentration uncertainty = sample mass uncertainty/sampler volume (for appropriate sampler channel).

3. Sample mass uncertainty = greater of
   a) lab blank uncertainty and
   b) square root of (s_analytical^2 + (M * sr_volume)^2).

   where:
       s_analytical = laboratory analysis uncertainty,
       M = sample mass
       sr_volume = volume uncertainty (relative) for appropriate sampler channel.

Notes:
1. Concentrations and concentration uncertainties are not defined for field and trip blanks because there are no sample volumes.
2. The sampler volume uncertainty (sr_volume) is currently assumed to be 5% for all sampler channels.

8.2.7 Generate Analytical Report Data in a draft database by running the program *New_Export_Report_Data*.

8.2.8 Check report data for duplicate rows by running the program *Check_for_Dups* with the argument of the latest batch number. Correct any problems and regenerate all steps to this part, if necessary.

8.2.9 Open the report generation program (Report.mdb) and verify that its tables are linked to the database copy used for quality assurance (QA) review. Re-link to the correct database with the linked table manager, if necessary.

8.2.10 Run the report program to generate draft copies of the output reports.

8.2.11 Notify the RTI QA Officer that the draft reports are ready for review.

8.2.12 After approval from the RTI QA Officer, repeat the procedures previously completed on the copy of database on the main database. Copy the final version of reports to the external Web site.

8.2.13 Remove any old reports from the Web site.

8.2.14 Notify the DOPOs that reports are ready.

8.2.15 Make CD copies of the Web site for the EPA Project Officer and for the RTI master project file.

8.2.16 Deliver CD to the EPA Project Officer.

9.0　Prepare results for AQS

9.1　Summary of Task

This procedure describes the preparation of the monthly AQS data report.

9.2　Procedure

9.2.1 Copy the current database over training database in preparation for the trial generation and review.

9.2.2 Run the *New_Fill_AIRS_Table* stored procedure with the argument of the batch to be delivered (this generates all data in the batch).

9.2.3 Using the main data entry application, set the event's delivery status to 5, AQS reprocessing needed, for all events that must be regenerated due to state reviews, etc.

9.2.4 Remove any data that needs to be reprocessed from the internal AQS processing table by running the *Delete_Challenged_Data* stored procedure with the argument of the AQS batch to be delivered.

9.2.5 Perform any needed recalculations by running the *New_Update_Challenged_Report_Data* stored procedure with the argument of the AQS batch to be delivered.

9.2.6 Check for any duplicate records by running the *Check_for_Dups* stored procedure with the argument of the batch to be delivered and fix any duplicates found.

9.2.7 Now that all records have been reprocessed, the AQS output file is generated by using the menu in the *AIRSApp.adp* program

9.2.8 Import the AQS file into the *AQSFIle* table in *AIRS_File_Checks.mdb* using the Import Text Wizard and choosing Delimited text format with vertical bar ("|") as the delimeter.

9.2.9 Run QA procedures, as described in Section 10, AQS QA Procedures.

9.2.10 Correct any errors and restart the process up to this step. Make note of any changes on the *QA review* form.

9.2.11 Send the file to the QA Officer to review the AQS data. Make any necessary corrections.

9.2.12 The QA Officer reviews and approves file.

9.2.13 Rerun procedures on main database.

9.2.14 Update the delivery status by running the *SetAIRSDeliveryStatus* program with argument of the AQS batch to be delivered.

9.2.15 Subdivide the approved AQS file into subfiles (typically two) to get files small enough for posting in the AQS.

9.2.16 Submit each sub-file to the AQS (see EPA's AQS documentation for procedures).

9.2.17 Notify the EPA WAM and Delivery Order Project Officers (DOPOs) that results have been posted.

10.0 AQS QA Procedures

10.1 Summary of Task

This procedure describes the steps needed to review and approve AQS data before posting them to the AQS.

10.2 Procedure

The overall procedure involves making a trial version of the AQS output file and reviewing the file for errors and inconsistencies (e.g., two results at the same site and date). The database is then corrected and a new AQS output file is generated and reviewed. This cycle is repeated until no errors are found in the AQS output file. The AQS output file is then posted to AQS. The procedure is described in more detail, below.

10.2.1 Run the AQS file. Make a copy of the current database for use in the QA review.

10.2.2   Generate a draft AQS file using the *AIRSApp.adp* database application. Make sure the application is connected to the database copy before file generation.

10.2.3   After the AQS file has been generated, import the file into the *AQSFile* table in the *AIRS_file_checks.mdb* QA software using the Import Text Wizard and choosing Delimited text format with vertical bar ("|") as the delimeter.

10.2.4   Check the data from the AQS file (now imported into the database) for errors by running the queries listed in the appropriate Quality Control Summary Sheet (Appendix A for routine samples and Appendix B for blanks). Query names for each check appear in italics on the *Quality Control Summary Sheet*.

10.2.5   Perform the first four steps in the *AQS Batch File Quality Control Summary Sheet* (Appendix A [or Appendix B for blanks]) to check for duplicates, modify records, new sites, and null start hours. If there are any new sites, check for complete monitor information. If the monitor information is incomplete, generate pdf files that show the monitor information for transmittal through the EPA DOPOs to request information from the appropriate agencies. After all monitor information is received, post monitor records in the AQS. Also add the AQS code, the POC, and the expected record count to the *Expected_AIRS_Counts* table in the *AIRS_file_checks.mdb* database. After the new sites are added to the expected counts table, the *Check_Record_Count2* query can be run to check the record counts. Also update the *Check_StartHour1* query batch number to the current batch number, and then run the *Check_StartHour3* query to check for odd start times.

Any errors found at this point must be corrected in the main database and marked for reprocessing. If events with duplicate dates cannot be resolved by reviewing the data in the database, send off a request to appropriate SHAL personnel so that they can review the hard copy paperwork and send inquiries to the site operator if necessary.

10.2.6   Subsequent Runs of AQS File. After all of the corrections have been made from the first run of the AQS file, an updated AQS file is prepared (by using the AIRSapp.adp program) and used to verify that all of the changes have been correctly made and to check that all of the state-requested changes have been correctly made.

To check the state-requested changes, run the *Temp_Trog13* query on the main database to get a list of all of the events in the current batch that are marked for reprocessing. Check in the appropriate batch folder for copied request change e-mails. Also check in the personal inbox for any other e-mails that may have requested that data changes be made. Make sure that all events marked for reprocessing have a requested change associated with them, as well as verify that

all e-mails with requested changes have the events marked for reprocessing in the database. If anything is missing, contact the person responsible for making the changes. If there is any question about the interpretation of changes, contact the QA Officer.

After all of events marked for reprocessing are verified, copy over the database and follow the procedures to generate an updated AQS file and import it into the *AIRS_file_checks.mdb* QA database. Run all of the queries that were previously run to verify that all errors were corrected. Because of the large datasets, the remaining queries are run on smaller static batch-specific tables, but these must be refreshed for each report batch. To perform this task, delete the data from the previous batch by running the *Delete_New_AIRS_Data* and *Delete_New_Report_Data* queries. Now refill the tables with the current batch's data by updating the batch number to the current batch number in the *Convert_Report_Data1* query, then, to append the new data, run the *Append_New_AIRS_Data* and *Append_New_Report_Data* queries. Now all of the queries specified on the Quality Control Summary Sheet (Appendix A) should be run to check the AQS file and to verify all of the changes. Note that the *Check_Delivered_Records1* query requires that the data from the previous AQS batch be appended to the *AIRS_Posted* table using the *Append_AIRS_Posted* query. This will verify that no events that were in a previous AQS file will appear in the AQS file.

If errors still remain in the AQS file, corrections should be made in the main database, and the AQS file should be reprocessed. Any manual changes to the AQS file, as well as any data withheld from the AQS file, should be noted in the Quality Control Summary Sheet.

After the final AQS file is generated, a completed Quality Control Summary Sheet, the printout results of all of the queries and printouts of all state-requested changes are submitted to the project's QA Officer for review. After the QA Officer has approved this quality control packet, the file can be loaded into EPA's AQS. Typically, the AQS file is split into 6 separate files of approximately 15,000 records to minimize loads on the AQS.

10.2.6 After all of the AQS data have been posted to AQS, run the AQS processing on the main database and mark events as delivered to AQS.

11.0 Database Backup

11.1 Summary of Task

Database backups are now provided by RTI's Information Technology (IT) server staff as part of their normal nightly back-up procedures. Procedures are covered in *SOP S0011-01: NT Server Data Backup and Storage of Backup Media*, which is available on RTI's internal Web site.

Appendix A

PM$_{2.5}$ Speciation AQS Batch File Processing Quality Control Summary Sheet

for Report Batch _____

Basic AQS File Checks (Required for AQS Update or by Format)

1. No duplicate rows in AQS file.

   Check_Duplicates1

2. No modify records that are not appropriate. Explain any re-deliveries of AQS events or manual modification to correct modifies.

   Check_Modifies1

3. Null data codes only occur when values are null; null data codes are null when values are not null. Also qualifiers do not occur without values or do not occur with null data codes.

   Check_Null_Value_Codes1

4. No sample values without Minimum Detection Limits (MDLs) and uncertainties (excluding the ambient temperature and barometric pressure parameters). No MDLs and uncertainties without sample values.

   Check_MDL_Uncertainty1

5. Added new locations or modifications of current locations. Please explain any changes.

   Check_New_Sites3

6. No records delivered in previous batches for the same date that are not modify records.

   Check_Delivered_Records1

7. Problems found during the AQS edits, which require manual edits in final the AQS file, please explain.

AQS File Checks Against Original Report Batch Data

1. Odd start times, if any, please explain.

   Check_StartHour3

2. Minimum and maximum sample dates for this AQS batch. If either date is not in the range for the current batch, please explain.

   Check_MinMax_SampleDate

3. Records in the AQS file that were not in original report, if any, please explain.

   Check_Not_in_Report1

4. Records in original report that are not in the AQS file, if any, please explain.

   Check_Not_in_AIRS1

5. Duplicates in original report (used only to rectify difference in counts).

   Check_Report_Duplicates

6. Rectify the number of records in the AQS file versus the original report to verify that all appropriate records in the report made it into the AQS file and that no spurious records were created in the AQS file that were not in the original report.

7. Record counts in the AQS file versus the expected record counts for each AQS code/POC. Explain any counts that do not match.

   Check_Record_Count2

8. Compare the values in the AQS file to those in the original report. Check 100% of those with an absolute difference of ≥ 1. Spot check those that are different by 10% or more, but with an absolute difference <1. Spot check any differences between 1% and 10% to ensure that differences are only caused by rounding. Explain any differences not caused by rounding.

   Check_Reprocessing

9. Compare null data codes in the AQS file to those in the original report. Explain any differences.

   Check_NullDataCode

10. Compare AQS validity codes in the AQS file to those in the original report. Explain any differences.

    Check_Validity

11. Compare MDL and uncertainty values in the AQS file to those in the original report. Check 100% of those that are different by 10% or more. Spot check any differences between 1% and 10% to ensure that differences are only caused by rounding. Explain any differences.

    Check_MDLs
    Check_Uncertainty

Verification of AQS Reprocessing Changes

1. Summarize changes made at the request of the states (check with person responsible for making the changes) and verify that these changes were updated with reprocessing.

2. Summarize changes made in response to problems found during the AQS processing and verify that these changes were updated with reprocessing.

Events in This Report Batch That Were Withheld from Update in the AQS

PM$_{2.5}$ Speciation AQS Batch File Processing Quality Control Summary Sheet

for Report Batch _____

Basic AQS File Checks (Required for AQS Update or by Format)

1. No duplicate rows in the AQS file.

   Check_Duplicates1

2. No modify records that are not appropriate. Explain any re-deliveries of AQS events or manual modification to correct modifies.

   Check_Modifies1

3. Null data codes only occur when values are null; null data codes are null when values are not null. Also qualifiers do not occur without values or do not occur with null data codes.

   Check_Null_Value_Codes1

4. No sample values without MDLs and uncertainties (excluding the ambient temperature and barometric pressure parameters). No MDLs and uncertainties without sample values.

   Check_MDL_Uncertainty1

5. Added new locations or modifications of current locations. Please explain any changes.

   Check_New_Sites3

6. No new records for site and date that were posted in a previous batch. However, update records (Action Code = U) for earlier events are acceptable.
   Check_Delivered_Records1

7. Problems found during the AQS edits that require manual edits in final AQS file, please explain.

AQS File Checks Against Original Report Batch Data

1. Odd start times, if any, please explain.

   Check_StartHour3

2. Minimum and maximum sample dates for this AQS batch. If either date is not in the range for the current batch, please explain.

   Check_MinMax_SampleDate

3. Records in the AQS file are not in original report, if any, please explain.

   Check_Not_in_Report1

4. Records in original report are not in the AQS file, if any, please explain.

   Check_Not_in_AIRS1

5. Duplicates in the original report (used only to rectify difference in counts).

   Check_Report_Duplicates

6. Rectify the number of records in the AQS file versus the original report to verify that all appropriate records in the report made it into the AQS file and no spurious records were created in the AQS file that were not in the original report.

7. Record counts in the AQS file versus the expected record counts for each AQS site code/POC. Explain any counts that do not match.

   Check_Record_Count2

8. Compare values in the AQS file to those in the original report. Check 100% of those with an absolute difference of ≥1. Spot check those that are different by 10% or more, but with an absolute difference <1. Spot check any differences between 1% and 10% to ensure that differences are only caused by rounding. Explain any differences not caused by rounding.

   Check_Reprocessing

9. Compare null data codes in the AQS file to those in the original report. Explain any differences.

   Check_NullDataCode

10. Compare AQS validity codes in the AQS file to those in the original report. Explain any differences.

   Check_Validity

11. Compare MDL and uncertainty values in the AQS file to those in the original report. Check 100% of those that are different by 10% or more. Spot check any differences between 1% and 10% to make sure that differences are only caused by rounding. Explain any differences.

Check_MDLs
Check_Uncertainty

Verification of AIRS Reprocessing Changes

1. Summarize changes made at the request of the states (check with person responsible for making the changes) and verify that these changes were updated with reprocessing.

2. Summarize changes made in response to problems found during the AQS processing and verify that these changes were updated with reprocessing.

Events in this Report Batch Withheld from AQS Update

Summarize any events in the current report batch that had to be withheld from updating in AQS (e.g., the state has not yet added the site to AQS, and thus we can not add monitors). Indicate what action will be needed to allow the withheld records to be posted in the future.