United States Environmental Protection Agency

Office of Research and Development

National Human Exposure Assessment Survey (NHEXAS)

Region 5 Study

Quality Systems and Implementation Plan for Human Exposure Assessment

Research Triangle Institute
Research Triangle Park, NC 27079

Cooperative Agreement CR 821902

Field Operations Protocol **RTI/ACS-AP-209-400**

Title: Procedure for Receipt, Processing, and Assembly of Analytical Results and Questionnaire Data Into a Database

Source: Research Triangle Institute

U.S. Environmental Protection Agency
Office of Research and Development
Human Exposure & Atmospheric Sciences Division
Human Exposure Research Branch

Notice: *The U.S. Environmental Protection Agency (EPA), through its Office of Research and Development (ORD), partially funded and collaborated in the research described here. This protocol is part of the Quality Systems Implementation Plan (QSIP) that was reviewed by the EPA and approved for use in this demonstration/scoping study. Mention of trade names or commercial products does not constitute endorsement or recommendation by EPA for use.*

TITLE: PROCEDURE FOR RECEIPT, PROCESSING, AND ASSEMBLY OF ANALYTICAL RESULTS AND QUESTIONNAIRE DATA INTO A DATABASE

SOURCE: Research Triangle Institute
Post Office Box 12194
Analytical and Chemical Sciences
Research Triangle Park, NC 27709-2194

AUTHOR(s):

_____ Date: 7/5/94

_____ Date: _____

_____ Date: _____

APPROVED BY:

Principal Investigator: _____ Date: 7/5/94

QA Officer: _____ Date: _____

STATUS:

IN PROGRESS: ☐

DRAFT: ☒

FINAL VERSION: ☐

REVISIONS:

No.	Date	No.	Date
0	‡	6	
1		7	
2		8	
3		9	
4		10	
5		11	

‡ Effective date of this version is the date of the last approval signature; revision 0 is the original version.

PROCEDURE FOR RECEIPT, PROCESSING, AND ASSEMBLY OF ANALYTICAL RESULTS AND QUESTIONNAIRE DATA INTO A DATABASE

TABLE OF CONTENTS

LIST OF FIGURES

1.0 SCOPE AND APPLICATION

The data management procedures described in this protocol are intended for application only during Phase I of the National Human Exposure Assessment Survey (NHEXAS). Assignment of samples to be collected from each study participant, the development of an "Information Shell", containing descriptive sample information, and the creation of a sample tracking scheme are addressed. This protocol also encompasses the receipt of analytical results from each of the participating laboratories, the introduction of the analytical data and questionnaire data into the database, and the extraction and application of descriptive data from the Information Shell. Finally, this protocol describes quality assurance and quality control procedures applied during data management.

2.0 SUMMARY OF THE METHOD

The principal function of the NHEXAS database is to receive, asssemble and process information from the participant query, sample collection and sample analysis components of the study. The Database Manager serves as the primary point of collection for all NHEXAS data. An Information Shell is constructed from pre-collection sample schedules, participant personal information and descriptive records compiled at the time of sample collection. Analytical data from sample analyses by consortium laboratories (RTI, EOHSI, Rutgers) is received as delimited ASCII files and is imported as a table either directly into Paradox for Windows or alternatively into Quattro Pro for Windows prior to import into the Paradox database. Sample analysis results from the federal laboratories (EPA, FDA, CDC) will be transformed from their initial format (ASCII, spreadsheet, word processing, hardcopy) into a Paradox-compatible format using an appropriate intermediate software tool. Questionnaire data will be manually entered into a data entry program and compiled into ASCII file formats.

3.0 HARDWARE AND SOFTWARE CONFIGURATION

3.1 Hardware

Upon receipt from the originating laboratory, all tabular components of the NHEXAS database will reside on DEC VAX 426 Mbyte hard disks and/or 90 Mbyte Bernoulli (IOMEGA Corp.) cartridges dedicated exclusively to NHEXAS. Application software will reside on an Everex 486DX2/50 personal computer running Microsoft Windows V 3.10. Complete nightly backups of all database files will be performed onto 8mm digital audio tape.

3.2 Software

Paradox for Windows (V 4.5; Borland International, Inc.) and Quattro Pro for Windows (V 5.0; Borland International, Inc.) application software packages will be used in the assembly and manipulation of the NHEXAS database. Pathworks (Digital Equipment Corp.) networking software will facilitate database file access on the of database files between the PC and the VAX hard disks.

4.0 CREATION AND APPLICATION OF THE INFORMATION SHELL

4.1 Assembly of Field Interviewer Pre-Sampling Data

A visual summary of the movement of pre-sampling data is presented in Figure 1. Pertinent pre-sampling information for each study participant will be recorded by the field interviewer and transmitted by FAX to the RTI Center for Survey Research (CSR) and, in turn, to the Database Manager in Analytical and Chemical Sciences (ACS). This pre-sampling material will include:

- ACS participant number
- CSR participant number
- participant's age
- participant's occupation
- participant's address
- participant's telephone number
- directions to participant's residence
- home type

- samples to be collected in this residence
- the amount of incentive to be dispensed to the participant

The field interviewer will transmit this information via FAX to CSR as soon as it is available so that . After recording information relevant to their activities, CSR will forward the pre-sampling data to the field supervisor, who will create a record of this information for each participant in the Information Shell database. This database will be created and augmented for each particpant using the relational database, Paradox (Version 4.5). Subsequently, the Information Shell will automatically produce a sampling schedule table for all samples to be collected from that participant. This sampling schedule table will be created both as hardcopy and as a delimited ASCII file. The file will be placed on a VAX disk which can be remotely accessed, via modem, by the field technicians. The accumulated sampling schedules, across all participants, in the developing Information Shell will be used by the database manager as the basis for determining the status of all samples.

4.2 Procedure

4.2.1 Receive the FAXed information from the Center for Survey Research and record the receipt into the database manager's notebook.

4.2.2 Manually enter the pre-sampling data as a record for each participant in the Paradox information shell.

4.3 Recording of Sample Collection Data

Sample collection data will be recorded in an automated Quattro Pro (Windows Version 5.0) spreadsheet at the time of sample deployment and recovery. The details of this operation are outlined in the protocol, RTI/ACS-AP-209-086, "Use of the Sample Collection and Custody Software".

4.4 Distribution of Collected/Exposed Samples to Analytical Laboratories

At the conclusion of sample collection for a given county, samples from that county will be distributed and shipped to the appropriate analysis laboratories (RTI/ACS-AP-209-___). To assist in this operation, a spreadsheet program will: 1.) summarize all of the samples collected, by sample type, for that county, 2.) create hardcopy chain-of-custody sheets to accompany the samples through their analytical journey, 3.) create a file for incorporation into the Information Shell which specifies the samples that have been sent for analysis and

the date of the shipment. This file will be transmitted by modem communication to the ACS VAX cluster.

5.0 RECEIPT AND PROCESSING OF ANALYTICAL DATA FROM RTI, EOHSI AND RUTGERS LABORATORIES

5.1 Overview

Figures 2 and 3 illustrate the movement of each type of analytical result from the originating instrument through transformation, assimilation of data from the information shell and calculation of individual analyte concentrations. Analytical technique, initial units, information shell contribution and final analyte units are indicated. Instrument data will be output in delimited ASCII format, whenever possible. Alternatively, instrumental output will be ASCII text. ASCII files will be intriduced to the Paradox database as individual tables and merged with appropriate data from the information shell, using the sample i.d. to produce concentration values. These concentration values will be created as calculated fields in each table.

5.2 Procedure

5.2.1 Receive the data from the originating instrument and record the receipt in the database manager's notebook, indicating the data format.

5.2.2 Import the data, as a table, into Paradox, and sort by sample i.d. number.

5.2.3 Define a field containing a mathematical formula for calculation of analyte concentrations.

5.2.4 Process the imported data table with the specified information from the information shell.

6.0 RECEIPT AND PROCESSING OF ANALYTICAL DATA FROM THE FEDERAL LABORATORIES

6.1 Overview

It is anticipated that analytical results will arrive from the federal laboratories in a diversity of formats, ranging from hardcopy to electronically-transmitted files. Figure 4 illustrates the movement of this data from its arrival at RTI through processing to analyte concentrations. The database manager will receive all analytical results from the federal

laboratories and determine the necessary database entry mechanism to be used to put the data into a Paradox database. Appropriate information from the information shell will be merged with the analytical data to calculate analyte concentrations.

6.2 Procedure

6.2.1 Receive the data from the federal laboratory and record the receipt in the database manager's notebook, indicating the data format.

6.2.2 For hardcopy data, manually enter the data into a Paradox table.

6.2.3 For data from a word processor, create a delimited ASCII file.

6.2.2 Import the data, as a table, into Paradox, using appropriate entry conversions and sort the table by sample i.d. number.

6.2.3 Define a field containing a mathematical formula for calculation of analyte concentrations.

6.2.4 Process the imported data table with the specified information from the information shell.

7.0 RECEIPT AND PROCESSING OF QUESTIONNAIRE DATA

Questionnaire data will be manually keyed by the RTI Center for Survey and Research Computing (CSARC) and provided to the database manager as ASCII files. This data will not be introduced to the database but will be retained on the VAX disks reserved for NHEXAS data.

8.0 RECEIPT AND PROCESSING OF SUPPORTING DATA FROM ANALYTICAL INSTRUMENTS AT THE CONSORTIUM LABORATORIES

Analytical results from quality control samples (blanks and controls) and preformance evaluation samples will be entered into the database as a table and identified by "batch" (e.g. analysis date). A table of supporting data will be created for each of the sample matrices/analyte types. Appropriate information, if any, from the information shell will be applied to the supporting data to calculate analyte concentrations.

9.0 DATABASE SECURITY

All access to the NHEXAS database components will be restricted to the database manager, the principal investigator and the quality assurance officer. Individual Paradox tables will be password protected and all files will reside on a VAX disk with access restricted to these staffpersons. Daily backups will be performed on the entire NHEXAS database and contributing components onto 8mm DAT tape. Backup tapes are locked in a fireproof safe remote to the source of the data.

10.0 QUALITY ASSURANCE

Quality assurance procedures will be conducted on the data at the point of import into the Paradox database but prior to assimilation of data from the information shell for all analytes and matrices. Comparisons will be made between the original data, as received from the source laboratory, and the Paradox table. Similarly, the calculated concentrations will be examined for correctness of assimilation of data from the information shell.

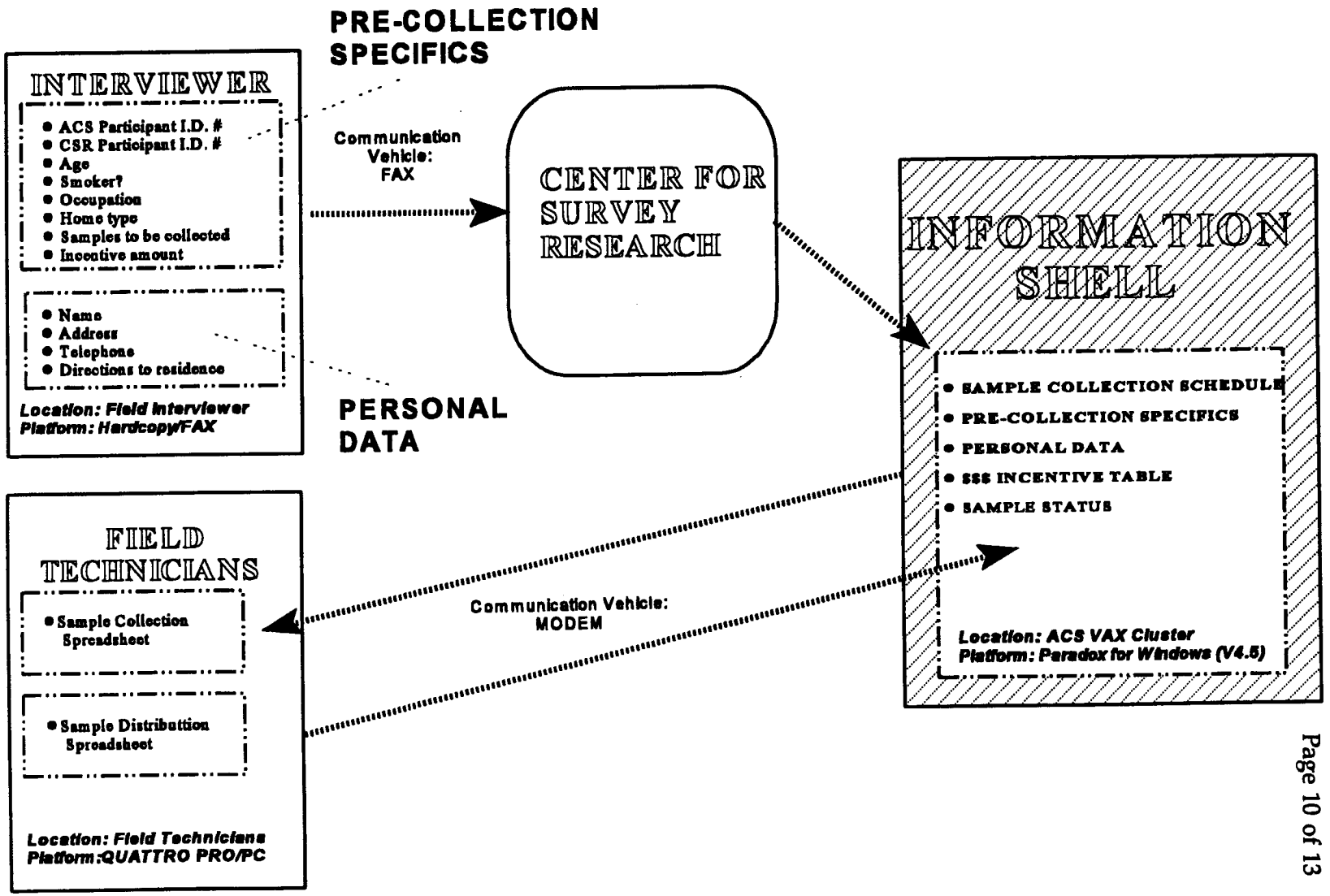

Figure 1. Creation and Application of the NHEXAS Information Shell

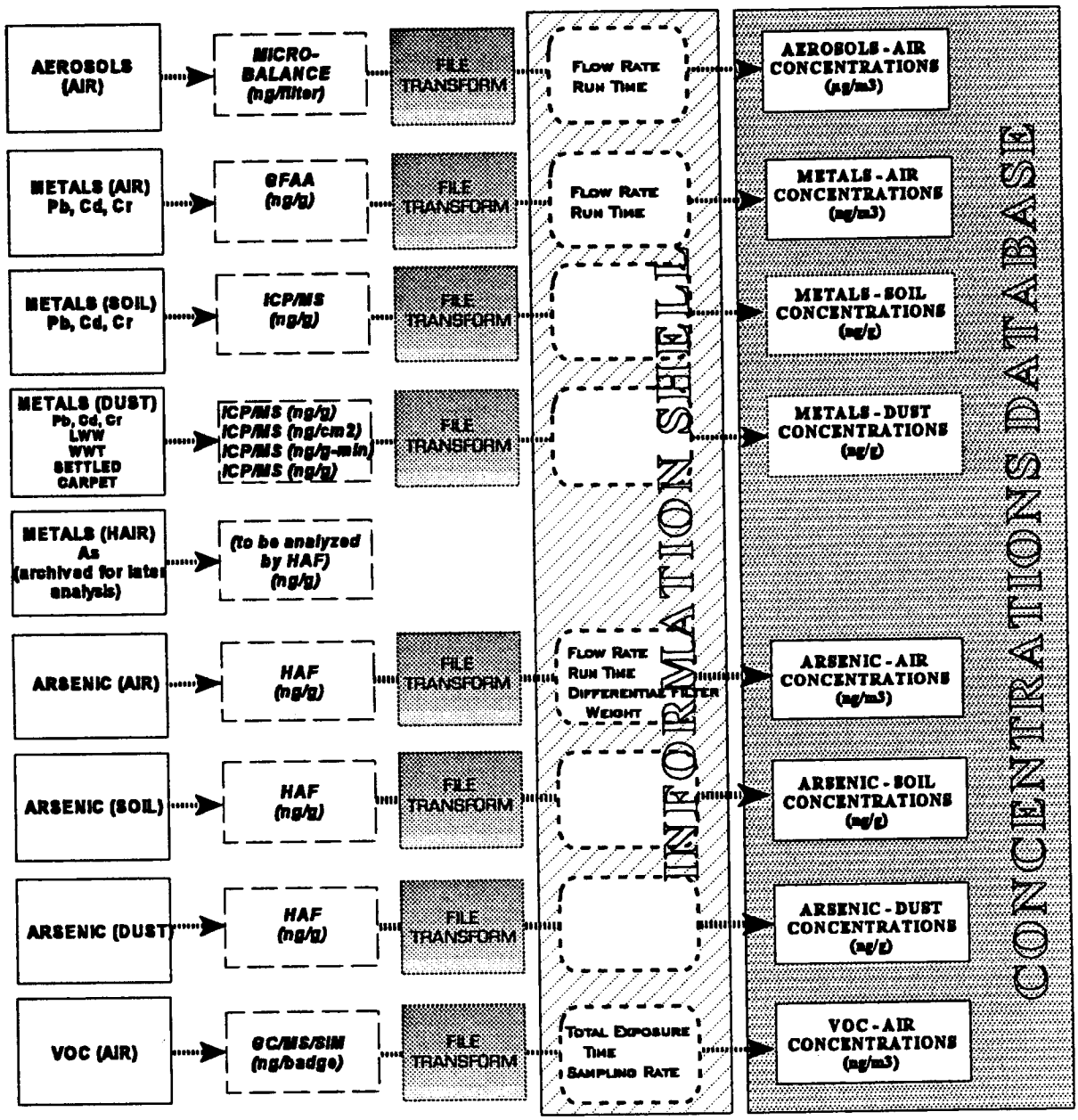

Figure 2. Receipt and Processing of VOC, Metals, and Arsenic Data from the Consortium Laboratories

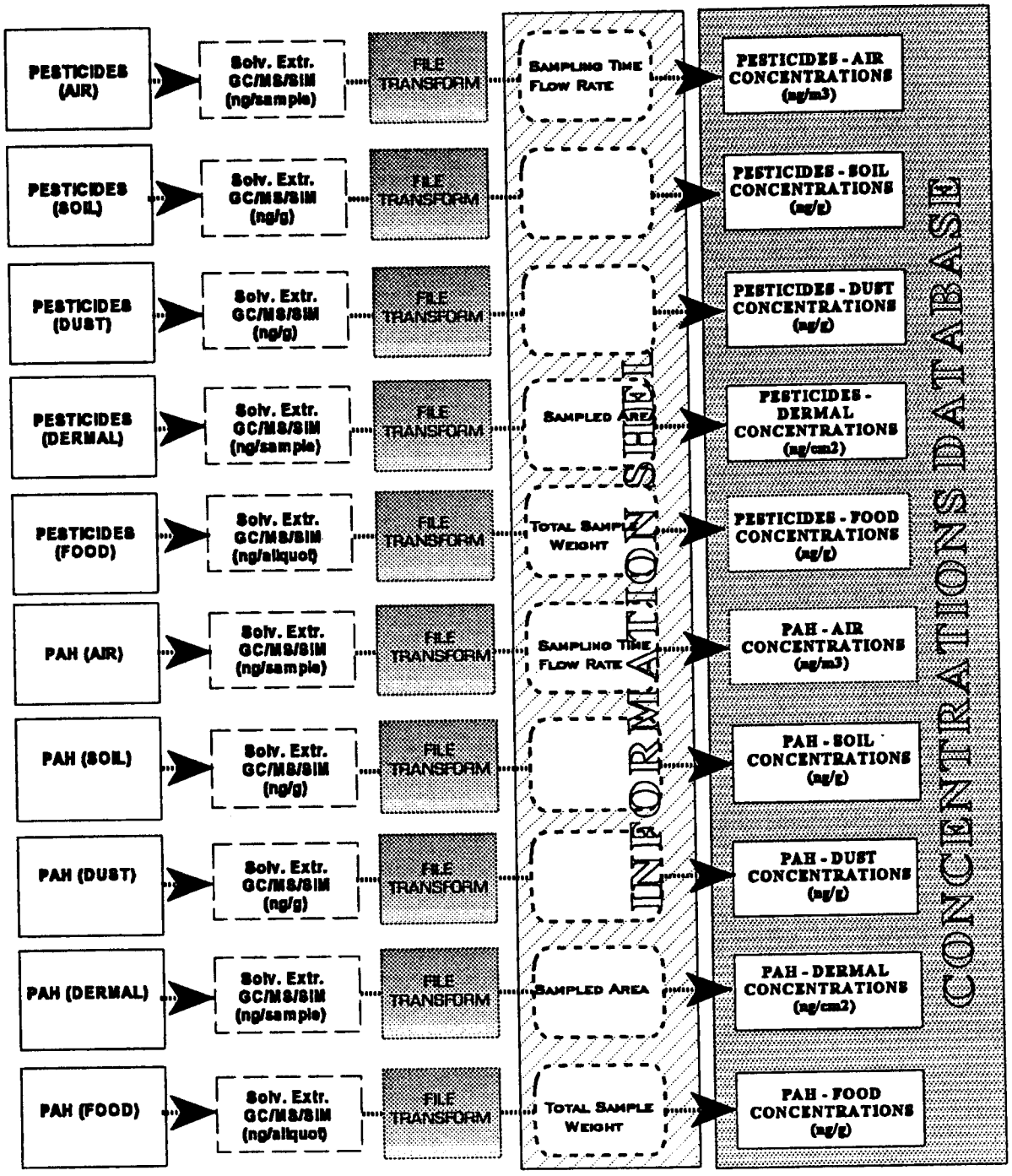

Figure 3. Receipt and Processing of Pesticides and PAH Data from the Consortium Laboratories

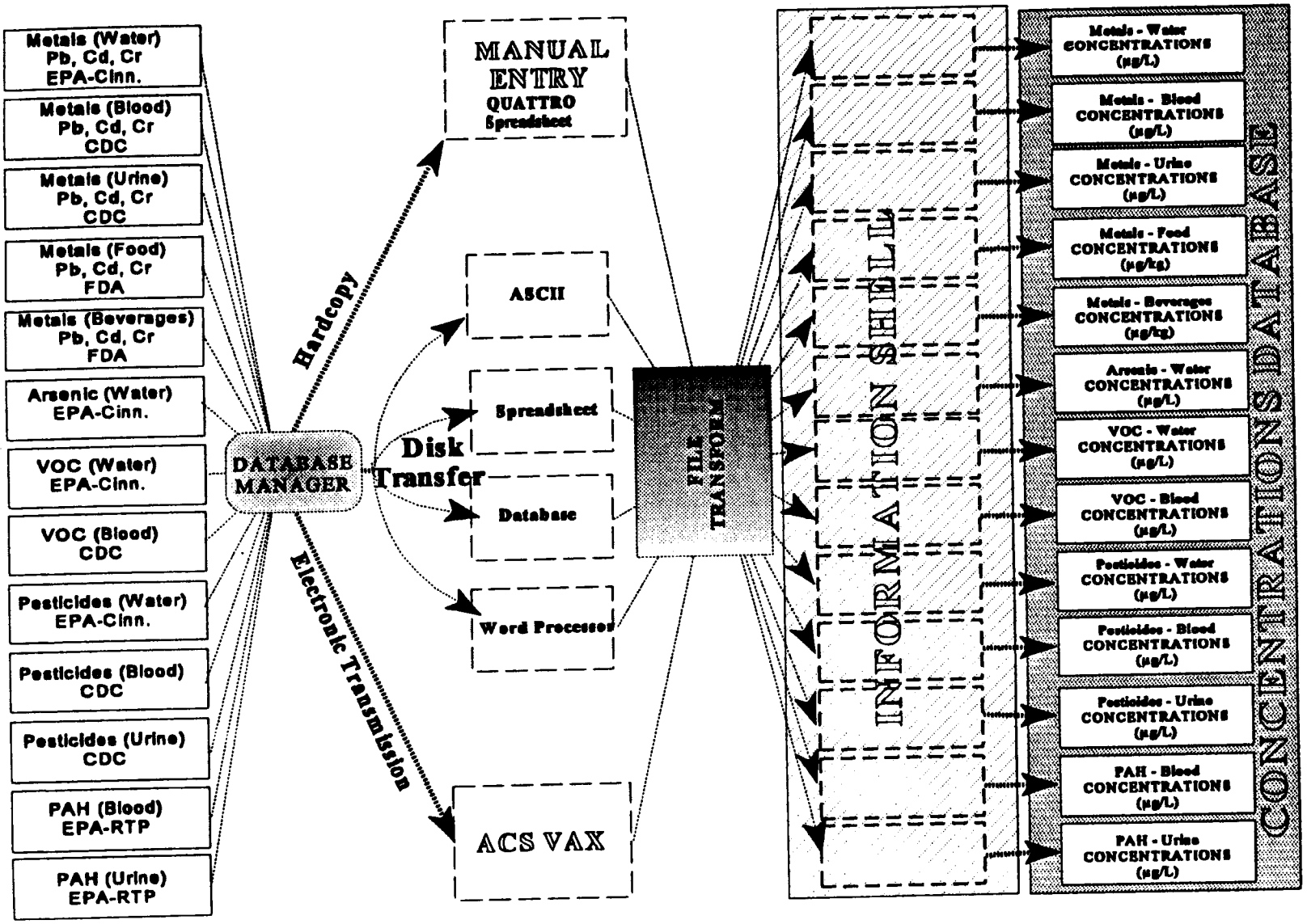

Figure 4. Receipt and Processing of Analytical Data from Federal Laboratories