

SECTION 3

QUALITY ASSURANCE APPROACH

Rigorous quality assurance (QA) procedures were applied to all stages of the survey to ensure that the data are of known quality. These procedures are detailed in two QA Plans¹; one covering the overall survey and one dealing specifically with tank testing. This section provides a brief overview.

In addition to the procedures outlined here and detailed in the QA plans, the survey was audited by EPA QA personnel. This involved site visits, productivity and response rate monitoring, and QA audits and reviews of home office processing procedures.

Data quality is assessed by its representativeness, completeness, accuracy, precision and comparability. The sample of establishments and their underground storage tanks must be representative of the population of establishments and tanks about which inferences are to be made. A complete, or almost complete, set of data must be collected; otherwise, its representativeness, accuracy and precision will be compromised. Accuracy (lack of bias) and precision must be sufficiently high so that confidence can be placed in the numerical value of the results. The methods employed must be well documented and allow

¹LUST Quality Control Procedures, National Survey of Underground Storage Tanks, Quality Assurance Plan, Westat, July 12, 1985; and National Survey of Underground Storage Tanks Draft Quality Assurance Program Plan for the Office of Toxic Substances, Midwest Research Institute, June 7, 1985.

results. The methods employed must be well documented and allow comparison of the results with other relevant studies. The succeeding pages outline procedures used at each stage of the study to assess and maximize data quality. Quantitative measures of data quality are presented in the appropriate sections throughout the report.

I. DEVELOPMENT OF SAMPLING FRAME AND SAMPLE SELECTION

The objective of the survey is to provide national and regional estimates of the number and proportion of underground motor fuel storage tanks that leak, and to investigate characteristics of leaking tanks. To achieve this objective, the information collected must be representative of storage tanks throughout the country. Representativeness was ensured by careful development of the sampling frames and methods of sample selection.

The entire contiguous United States was divided into Primary Sampling Units (PSUs) consisting of counties or groups of counties with a minimum number of gas stations and other fuel-related business establishments. A probability sample of PSUs was drawn, and sample summary statistics were compared with summary statistics calculated from the entire frame to check that the random selection procedure had been applied correctly.

Section 4 and Appendix A detail the construction of frames and sample selection for the fuel establishment, large establishment and farm samples. In each case, checks were made between frame and sample summary statistics.

II. PREPARATION OF QUESTIONNAIRES AND INTERVIEWS

Versions of the questionnaire and inventory recording forms were pretested, both informally and in extensive formal pretests. The questionnaire, together with an instruction booklet, was sent to the respondent prior to the interview to allow the respondent to look up information not immediately available and thereby improve the quality (accuracy and completeness) of the response. Questionnaire packages were sent by certified mail to provide a record of receipt.

Interviewers were carefully selected and underwent a five-day training session that included instruction at both Westat and MRI in Kansas City. They were provided with a detailed interviewer's manual that explained project and administrative procedures, and provided item-by-item specifications for the questionnaire.

III. INTERVIEWING

One or more interviewers were assigned to each PSU. They received call records and labels for each of their assigned cases, and were to complete a questionnaire for each eligible case. All survey materials for completed work were sent to the home office weekly. Call records for ineligible as well as eligible cases had to be returned with the reason for ineligibility clearly marked on the call record. Interviewers were required to make weekly progress reports by telephone.

In addition to questionnaire administration, interviewers were responsible for conducting a meter accuracy test, reviewing inventory procedures with the respondent, and collecting other information on the location and nature of the establishment. The measured meter errors were used to adjust the meter sales figures to give actual quantities dispensed each day. This, in turn, reduces the incidence of false-positive results from inventory reconciliation analyses when quantity dispensed is compared to stick readings. Review of inventory procedures was aimed at improving the respondent's ability to understand and correctly complete the inventory recording forms.

Interviews from the field were checked for completeness, given a final interview status code, and logged into an automated receipt control system. Refusals to comply with this mandatory survey were reported directly to EPA's Office of Enforcement and Compliance Monitoring. Most refusals eventually complied, and the overall response rate was 99.3 percent. Approximately 78 percent of the eligible establishments furnished complete or partial inventory data. Difficulties in gathering inventory data are discussed in Section 5.

Initially, replacements for ineligible (i.e., out of scope because no below-ground tanks, no motor fuel, etc.) fuel establishments were selected after their ineligible status had been determined. An unexpectedly high ineligibility rate of about 50 percent made this procedure impractical. Subsequently, the lists were oversampled, and the interviewers screened for eligibility in the field. This made replacement errors less likely.

IV. DATA TRACKING

An automated receipt control system was developed to maintain a record of the status of each data type for each establishment. The data types and activities logged into the system include:

- o Certified mail cards - date received;
- o No tank certifications - date received;
- o Interview status - Complete, Partially Complete, Ineligible, etc., with appropriate dates;
- o Contact name, mailing address and telephone number;
- o Prompt call status - date of call to prompt for inventory records;
- o Inventory status - Complete, Partially Complete, Refused, etc.;
- o Tightness test flag - establishments selected for tightness testing;
- o Tightness test status - Complete, Partially Complete, Refused, etc.;
- o Map status - Availability of soil and hydrogeologic data; and
- o Final status - Computed from the above status fields.

The system enabled the survey to be monitored on a continuous basis by providing document control and producing various lists and reports. Discrepancy reports indicating missing data types, such as missing inventories or missing "No

Tank" certification, provided measures of the completeness of the survey.

V. SELECTION OF ESTABLISHMENTS FOR TIGHTNESS TESTING

The frame was developed from the information in the receipt control system and checked against manual records before drawing the sample. Weighted and unweighted summaries of the sample were calculated to check that the sample of 198 establishments was drawn correctly from the 876 eligible business and government establishments. All 20 farms were scheduled for testing.

Notification packages were sent to each selected establishment with a request for a return receipt. Four to seven days after the mailing, each respondent was called to verify that the package had arrived and that the testing procedure was understood.

VI. TIGHTNESS TESTING

A. Training Tightness Teams

To assure that testing conducted by the several tank testing crews would produce accurate and reliable results, a formal training program was conducted prior to the beginning of the testing program. All of the crew members participating in the training were employees of O.H. Materials, Inc. or Double-Check,

Inc. and had been trained and certified by Heath Consultants (the manufacturer of the testing equipment) prior to the training for the national survey. The two main objectives of the training were to assure that all crews would follow the same test procedures, and to provide advanced tank testing training.

The MRI field data analysts who accompanied each tank testing crew received a four day training session that covered procedures for collecting data, preparation of site diagrams, computer operations, and data transmission. The analyst's training sessions included hands-on practice experiences in data entry, data transmission and recordkeeping.

It was necessary to add six additional test teams at approximately halfway through the testing program. The test crews were subcontracted through Protanic, Inc, and were selected based on their past performance records, and years of experience. The added crews were trained in the testing protocols by on-site trainers or by sessions held via telephone, and the test crews were teamed with experienced MRI data analysts who monitored the testing procedures. All data collected by the added test teams received extra verification review by Protanic and MRI staff prior to entering the data in the analysis files.

B. Instrument Calibration

The Petro-Tite equipment used to collect the leak rate data includes a thermistor in a probe that is inserted into the fuel tank to monitor fuel temperature and a thermal sensor box which provides an absolute temperature measurement within three degrees

Fahrenheit. The electronic circuit in the thermal sensor box was checked each time the equipment was set up (on site) to assure that the circuit detected a simulated temperature change to within approximately 0.003 degrees Fahrenheit, before the equipment was used in the test.

Thermistors, glass thermometers and barometers used to collect ancillary environmental information such as air temperature, barometric pressure, and surface and subsurface temperature were calibrated at MRI prior to field use, and the calibration data were entered as a part of the project file.

C. Field Inspections

Site visits to tank test operations were conducted by the MRI project leader, the MRI quality assurance coordinator, the field coordinator, and EPA staff. The purposes of the site visits were to evaluate the test protocols, and to assess the performance of the field crews.

D. Data Management and Analysis

A computerized receipt control file was used to track the test data. As the data were processed and reported, that stages of the process were entered into the receipt control file. Weekly reports were printed for project management and reporting.

The tank test data were entered into a portable computer in the field, using a thoroughly tested data entry program. These data files were transmitted to MRI by telephone, where they were stored on disks. The hard copy of the data on the original Petro-Tite sheets, the field copies of the disk files, and the data collected on ambient conditions was sent to MRI on a weekly basis. The field disk files were compared to the transmitted disk files using a computer utility program, in order to detect any transmission errors. All of the data elements to be used in subsequent leak analyses were printed out and hand checked against hard copy.

The data file for each test was hand checked for measurement problems and outliers, which were eliminated prior to the analysis of the file to determine test status and leak rate. Each tank test analysis was checked individually to ensure that the formulas for calculating the leak rate and standard error were correct. Leak rate, test status, and environmental data were abstracted from hard copy and computer files, and coded, keyed, key-verified, and edited prior to being merged with the questionnaire analysis file.

E. The Retest Program and Results

Three types of retesting were carried out to investigate different sources of variation in the test results. Back-to-back retests (on the same day and by the same crew) were conducted to investigate the stability of test results over time. A leak simulation retest was used to check the accuracy of the test in a situation with a constant leak of known rate. "Complete retests"

were independent tank tests usually conducted on different days with different crews from the original test, and involving rescheduling the test and refilling the tank.

A detailed analysis and discussion of the retest data is reported in Section 7 and Appendix D of this report.

VII. DATA HANDLING AND MANAGEMENT

Questionnaires were batched in groups of ten and tracked through the entire survey processing operation. Rules for coding and editing were published in a Coding Manual that was updated periodically to include new codes. Coding problems were referred to a supervisor for resolution. Inventory records were reviewed prior to coding so that missing or incorrect information could be recovered or corrected by a phone call to the respondent as soon after receipt as possible. Approximately 90 percent of respondents received prompt or data retrieval calls in an effort to obtain complete and accurate data on both the questionnaires and inventory records. All coded questionnaires were verified by the coding verifier.

All data were key-punched and then key-verified by a second operator and transferred by telephone link to EPA's NCC computer facility at Research Triangle Park in North Carolina. The data were machine edited to check that each data element is in range, that skip patterns were correctly followed, and that answers to related questions are consistent. Errors were corrected and the edit program rerun until no more problems were found. Frequency distributions and various tables were generated to check on data

quality. These allowed outliers to be identified and checked to ensure that they are "true" values.

Careful coding and data entry procedures ensure that the final data files accurately represent the information collected by the survey.

VIII. DATA ANALYSIS

National estimates of number of establishments with tanks, number of tanks, percentage of leaking tanks and leak rates were obtained using the methods described in succeeding sections. Where appropriate, the estimates are accompanied by confidence bounds or standard errors to indicate precision. Confidence bounds are narrowest (i.e., the estimates are most precise) for estimates based on data from the entire survey. Estimates based on subsets of the data (e.g., data from individual regions or for tanks of a particular age) have broader confidence intervals.