VERIFYING HISTORICAL DATA

Data used in a scientific study must be evaluated for:

RELEVANCE: Are the data and information relevant to the study?

ACCURACY: Are the data and information accurate?

CONSISTENCY: Are the data and information consistent and reproducible?

RELIABILITY: Can the sources of data and information be located?

SUFFICIENCY: Are the data and information sufficient to reach a conclusion?

Is there complete and reliable information on Rocky Flats releases?

One of the most difficult problems in analyzing releases of contaminants that occurred in the past is obtaining complete and reliable information. Researchers must reconstruct events that happened many years ago using whatever information is available. In the study of releases from the former Rocky Flats Nuclear Weapons Plant and the resulting public health impacts (Historical Public Exposures Studies on Rocky Flats), several issues affected the availability and reliability of the data:

- The research covered nearly 40 years starting in 1952, when records were maintained manually. Although record-keeping practices have changed significantly in the past four decades, these old records provided important information about past operations.
- Some of the records and data from the early years of plant operation were missing or incomplete. Reconstructing accidental releases, such as plutonium released in the 1957 fire, was especially difficult because questions remained about the amount of plutonium that was in the building before and after the fire. Similar data gaps existed for other incidents.
- The U.S. Department of Energy (DOE), has begun to declassify some information that was previously restricted for national security reasons. In all dose reconstruction studies related to the Nuclear Weapons Complex, such as the Rocky Flats study, access to needed records caused some delays due to classification issues. The Health Advisory Panel of the Historical Public Exposures Studies on Rocky Flats requested that DOE declassify additional data needed for the studies. Information of a national security nature, such as weapons design data, is never needed in dose reconstruction.
- Because of the secrecy that surrounded nuclear weapons production at Rocky Flats, especially in its first two decades of operation, some individuals doubted the reliability of historical information from the plant.

What were the sources and types of data needed for dose reconstruction?

Researchers used several different sources of information and types of data to reconstruct routine and accidental releases of contaminants from Rocky Flats. Sources of information about Rocky Flats activities included the DOE, plant operators, local, state and other federal agencies, current and former employees and others with an interest in the plant. Original and basic site data were more useful and credible for reconstructing past releases from the site than records that summarized the information, such as monthly or annual reports.

Specific information sources included written logbooks and other documentation related to plant operations and data from emissions monitoring and environmental sampling of air, soil and water. Oral accounts from knowledgeable sources were used to augment written information and data.

Regardless of the information source or type of data, researchers must always critically examine the information for accuracy and credibility. A methodical approach is the best means of evaluating data and

information to be used in a scientific study. Several questions need to be answered for the data evaluation:

Are the data and information RELEVANT to the study?

The Rocky Flats dose reconstruction study (Historical Public Exposures Studies) was designed to build a record of past releases of contaminants that occurred either during routine plant operations or as a result of accidents. Data and information used in the reconstruction therefore needed to be relevant to the particular release event.

Are the data and information ACCURATE?

Accuracy pertains mostly to written analyses or numerical data. To judge data accuracy, researchers must evaluate whether the method used to generate the data was adequately documented and appropriate.

For example, publicly available reports on Rocky Flats operations were compiled as summaries of data for specific time periods. The accuracy of this summary information was evaluated by going back to the original raw data recorded daily by the technician or operator. The researcher looked for consistency between the raw data and the summary report and often performed an independent calculation to check the accuracy of the conclusions.

Are the data and information RELIABLE?

To check the quality and reliability of data, the original source of that data was located when possible. Eyewitness accounts also helped locate written documentation. Information which could not be validated by recorded data or eyewitness accounts required further scrutiny to assure its relevance, accuracy and consistency.

Are the data and information CONSISTENT and REPRODUCIBLE?

In some cases, more than one source or type of data was available to reconstruct some of the past contaminant releases from Rocky Flats. Data from different sources, such as air monitoring done by a regulatory agency and that done by the facility itself, were compared for consistency.

Using multiple sources and types of data for an analysis usually improves the quality and reliability of the results. If the data do not agree, further research may be needed to explain and document the reasons for the differences. In addition, different methods or data sources can be used in the analysis to see if the results can be reproduced. For example, the amount of a certain contaminant predicted to be in the environment using a mathematical model can be compared with actual measurements of the contaminant detected in environmental samples. If the results are similar, the model can be used to make predictions when actual measurements are not available.

Are the data and information SUFFICIENT to reach a conclusion?

If past events were well-documented and the available data are reliable and sufficient for dose reconstruction, conclusions can be made with confidence. Sometimes data and information are incomplete, but an adequate basis exists for making assumptions. In this case, the assumptions must be clearly stated and the resulting limitations must be documented.

For instance, in estimating releases to the atmosphere, the actual wind and weather conditions at the time and place of the release may not be available. In this case, data from the same season in several other years can be used in mathematical models to simulate how and where the contaminants were transported under typical wind and weather conditions.

Scientists can sometimes use alternative approaches if the available data are not adequate for an analysis. For example, if air emissions data are not complete, one approach would be to estimate the release using information from a similar process at another site, using a model or laboratory experiment

to simulate the process that produces the emissions (such as burning materials under specific conditions). The available data, though limited, can then be used to check the model.

Summary

Reconstructing past events, particularly those that took place decades ago, is extremely challenging. In scientific research such as the Historical Public Exposures Studies on Rocky Flats, the objective was to provide the best possible answers given the limitations of available information and data, considering their relevance, accuracy, consistency, reliability and sufficiency.