Pacific Northwest
NATIONAL LABORATORY

PNNL-32733

Recommendations for Radiological Data Assessment Implementation

March 2022

Eric M Becker
Angela S Moore
Jonathan B Napier
Nicole R Benker
Kenneth M Thomas

Prepared for the U.S. Department of Homeland Security
contract 70RSAT21KPM000055, under a Work for Others Agreement with the
U.S. Department of Energy under Contract DE-AC05-76RL01830

DISCLAIMER

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

PNNL-32733

Recommendations for Radiological Data Assessment Implementation

March 2022

Eric M Becker
Angela S Moore
Jonathan B Napier
Nicole R Benker
Kenneth M Thomas

Prepared for the U.S. Department of Homeland Security contract 70RSAT21KPM000055, under a Work for Others Agreement with the U.S. Department of Energy under Contract DE-AC05-76RL01830

Pacific Northwest National Laboratory
Richland, Washington 99354

This page intentionally left blank.

Summary

Fast, agile, and timely decisions depend on having the correct information at the right time. When the stakes are high during a nuclear or radiological event, uncertainty in the information presented can overwhelm the most seasoned response leadership and delay making the decisions that matter (e.g., shelter versus evacuation, medical countermeasures, resource deployment). Data quality practices add confidence and defensibility to decisions and courses of action. Whether working with agriculture products or families, officials with high-quality data make better decisions concerning relaxing or prolonging protective actions. Data quality practices also provide a sound technical basis for a decision if it is challenged.

Since thorough verification, validation, and data quality assessment (DQA) processes may delay incident commanders and elected and appointed officials from making key decisions, the U.S. Department of Homeland Security (DHS) National Urban Security Technology Laboratory (NUSTL) tasked Pacific Northwest National Laboratory (PNNL) to develop tools and guidance for the federal, state, local, tribal, and territorial (FSLTT) responders using research, discussions with select FSLTT responders, and the experiences of subject matter experts in incident response. This report provides NUSTL with recommendations on best practices for verification, validation, and data quality assessment for the data collected by responders during a radiological or nuclear event. The purpose of this report is to inform the development of a DQA toolkit aimed at the needs of data assessors during the response to a radiological incident.

This report describes the entire data life cycle, the process for applying data quality practices, as a foundation upon which the methods and practices change the data into defensible information from which commanders make their decisions.

Three data quality principles guide all data life cycle practices:

1. Measure data quality relative to the question the data answers.
2. The data quality process provides a set of conclusions implied by the data, and it does not provide a set of prescribed decisions that a commander must make.
3. The data quality process is iterative.

Data assessor is the term for the individuals who apply verification, validation, and DQA practices, collectively referred to as the assessment phase of the data life cycle. A data assessor verifies that all critical information has been collected and reported, validates that the measured data meet the requirements of the response objective, and draws conclusions from the available information by considering any issues or anomalies with the provided data and drawing conclusions about the data. The limitations in time and resources for radiological incident response mean that the assessment phase process must be adjusted for incident response. Therefore, a graded approach should be taken where the data assessment process starts with basic checks for data quality in the early phase and progresses to more rigorous checks and statistical tests during the intermediate phase.

The research performed has led the team to make seven recommendations for the DQA toolkit that will be the focus of Task 3 of the project:

1. An ideal DQA toolkit will include

- a guidance document for how to apply the data assessment process to common radiological incident response objectives, and

- an interactive electronic data assessment workbook that implements the data assessment methods described in the guidance document.

2. The DQA toolkit guidance and workbook should not require data assessors to reference any other data quality guidance and should not require any computational tools more complex than a pocket calculator or smartphone.

3. The methods implementation of the DQA toolkit should be organized around response objectives to reinforce the idea that each objective is different and should be addressed individually.

4. Statistical methods in the DQA toolkit should be restricted to intermediate phase objectives, and not be used for early phase objectives due to the greater urgency of early (emergency) phase actions.

5. The DQA toolkit should only consider the early and intermediate phase. It is expected that, by the time of the late (recovery and remediation) phase, the full in-depth assessment phase process should be practical.

6. The DQA toolkit should be primarily targeted toward SLTT personnel who need to analyze radiological measurement data and present conclusions about the available information to incident response leadership, such as section leaders, incident commanders, and elected and appointed officials.

7. The DQA toolkit workbook should be based on existing software of verified quality.

The team also documented additional recommendations in this report that will help improve the data assessment process, but which are outside the scope of the DQA toolkit effort.

Acknowledgments

The authors would like to thank the following individuals for their insights and recommendations in developing this work:

- Washington State Department of Health Office of Radiation Protection, Radiological Emergency Preparedness Section – Mark Henry, Section Manager; Jack Wiley, Radiation Health Physicist
- New York City Department of Health and Mental Hygiene – Dr. Mark Maiello, Radiological Projects Planning Manager; Rachel Hunicke, Emergency Response Planner (Ms. Hunicke is now with the National Urban Security Technology Laboratory)
- Fire Department of New York –Timothy Rice, Battalion Chief, Hazardous Materials Operations & Weapons of Mass Destruction Branch Director
- Pennsylvania Department of Environmental Protection – David J Allard, Director of Bureau of Radiation Protection
- U.S. Department of Energy, National Nuclear Security Administration – Jeffrey Chapman, Technical Advisor
- U.S. Environmental Protection Agency, Office of Radiation and Indoor Air – Jennifer Mosser, Quality Assurance Manager & Science Advisor
- Federal Emergency Management Agency, Chemical, Biological, Radiological, and Nuclear Branch Office – Jonathan Gill, Nuclear Incident Response Team and Nuclear Radiological Incident Task Force Program Manager.

The authors would also like to thank Andrew Karam and Orly Amir at the U.S. Department of Homeland Security National Urban Security Technology Laboratory for sponsoring this work and for their excellent feedback and guidance.

Acronyms and Abbreviations

CBRN	chemical, biological, radiological, and nuclear
cpm	counts per minute
DHS	U.S. Department of Homeland Security
DOE	U.S. Department of Energy
DQA	data quality assessment
DQO	data quality objective
EOCs	emergency operations centers
FEMA	Federal Emergency Management Agency
FSLTT	federal, state, local, tribal, and territorial
GM	Geiger-Mueller
MQOs	measurement quality objectives
NNSA	National Nuclear Security Administration
NUSTL	National Urban Security Technology Laboratory
PNNL	Pacific Northwest National Laboratory
PRD	personal radiation detector
QA	quality assurance
QC	quality control
RDD	radiological dispersion device
SLTT	state, local, tribal, and territorial
UTL	upper tolerance limit
V&V	verification and validation

Contents

Figures

Tables

1.0 Introduction

Fast, agile, and timely decisions depend on having the correct information at the right time. When the stakes are high during a nuclear or radiological event, uncertainty in the information presented can overwhelm the most seasoned response leadership and delay making the decisions that matter (e.g., shelter versus evacuation, medical countermeasures, resource deployment). Data quality practices add confidence and defensibility to decisions and courses of action. Whether working with agriculture products or families, officials with high-quality data make better decisions concerning relaxing or prolonging protective actions. Data quality practices also provide a sound technical basis for a decision if it is challenged.

Since thorough verification, validation, and data quality assessment processes may delay incident commanders and elected and appointed officials making key decisions, the U.S. Department of Homeland Security (DHS) National Urban Security Technology Laboratory (NUSTL) tasked Pacific Northwest National Laboratory (PNNL) to develop tools and guidance for the federal, state, local, tribal, and territorial (FSLTT) responders using research, discussions with select FSLTT responders, and the experiences of subject matter experts in incident response. This report provides NUSTL with recommendations on best practices for verification and validation (V&V), and data quality assessment (DQA) for the data collected by responders during a radiological or nuclear event.

This report summarizes V&V and DQA methods and procedures. It also provides details on the assumptions and reasoning used by the PNNL team to select the best incident response V&V and DQA practices and discusses how these relate to the information gathered by FSLTT and other stakeholders. The appendices to this report augment information provided in the main text. Appendix A contains a glossary of terms, Appendix B provides an example of the progression of data assessment during incident response recommended by this report, and Appendix C provides an example of a data collection form that has incomplete responses and discusses the ramifications of incorrect or partial information.

Applying these best practices improves the timeliness and accuracy of official data products by streamlining the data quality process while producing more data of confirmed sufficient quality for decision-making. This report informs the development of a toolkit aimed at the needs of data assessors during the response to a radiological incident. The toolkit will be developed in the next phase of this project.

Beneficiaries of best practices in verification, validation, and data quality assessment

Data assessors may include a variety of persons filling any of several roles within an emergency operation center, or special advisors to elected and appointed government officials concerning radiological hazards during an incident. For example, an incident commander or state official may assign a Radiological Operations Support Specialist (FEMA 2022) the responsibility for data assessment. For the purposes of this report, data analysis includes data assessment, as well as operations like converting between measurement units, calculating derived values, and other data manipulation tasks that fall outside of data assessment.

In this report, the term **response leadership** refers to section leaders, incident commanders, elected and appointed officials. Each of these roles will make decisions during a radiological incident. Section leaders and incident commanders may make decisions about how to direct the response. Elected and appointed officials at all levels of government may decide on and order the implementation of protective action decisions. Data are often requested to support both kinds of decisions, and the data assessor provides one of the many inputs response leadership receive to make their decisions. The decision-maker may weigh several other considerations (sometimes conflicting), such as the dose that emergency responders may receive versus the urgency of the decision that will be supported, or the potential effects of an evacuation versus the dose the population is projected to receive.

2.0 The Data Life Cycle

The data life cycle determines if measurement data are traceable, defensible, and—where applicable—technically valid. The data life cycle provides a firm technical basis for decisions if they are challenged. The data life cycle is a process that includes three phases: planning, implementation, and assessment, which occur before, during, and after data collection, respectively.

The data life cycle comprises three principles.

There is no absolute measure of data quality; it is always measured relative to the question the data answers.

Put another way, "data quality, as a concept, is meaningful only when it relates to the intended use of the data" (EPA 2000). This principle means that there is no absolute measure of good and bad data quality, and that data quality must be assessed independently each time the measurement values are used. One example of this principle is to consider a set of measurements made with a Geiger-Mueller (GM) counter, provided in counts per minute (cpm). These measurements might be considered as having sufficient quality to determine whether any radioactive contamination is present but would not be sufficient for identifying or quantifying specific radioisotopes present in the contamination. Some questions may even seem to be asking the same thing but are not. For example, consider a set of GM counter measurements that are made in a particular neighborhood and then are used to answer a question about that neighborhood. If the same question arises, but is focused on a different neighborhood, those measurements are likely not appropriate and new measurements may need to be collected.

The data life cycle provides a set of conclusions about the available information; it does not provide a set of prescribed decisions that a commander or elected or appointed official must make.

Throughout the data life cycle, potential conclusions that may be drawn based on the information provided are identified. Some conclusions may also include potential actions that may be taken based on the question asked. However, these should always be considered actions that *might* be taken and not actions or choices that *must* be made. There are always other factors that affect decision-making, such as resource availability, cost, human factors, and hazards other than radiation. Therefore, the output of the data life cycle is just one input to a decision and must be considered in combination with other information and factors. For example, radiological measurement values might support a decision to reopen a city park for recreation, while other factors (e.g., the presence of asbestos) might support a decision that the park remain closed.

The data life cycle is iterative; some steps in the process may be repeated depending on the information collected and the conclusions provided.

Although the phases and steps of the data life cycle and the assessment phase, are presented in a specific order, the outcome of one step will sometimes be to return to an earlier step, or even a step in an earlier phase. For example, during the assessment phase, one may discover that many of the measurements collected are not of sufficient quality. In such cases, one may need to revisit the data quality objective (DQO) and quality assurance (QA)/quality control (QC) steps to establish the type, accuracy and amount of relevant information needed and then make sure that the data are collected and reported with sufficient quality. Figure 1 illustrates the data

life cycle and provides an expanded review of the three steps in the assessment phase, which are described in the following sections. Since the guidance that will be developed on this project is limited in scope to the assessment phase, the authors provide only brief descriptions of the planning and implementation phases of the data life cycle.

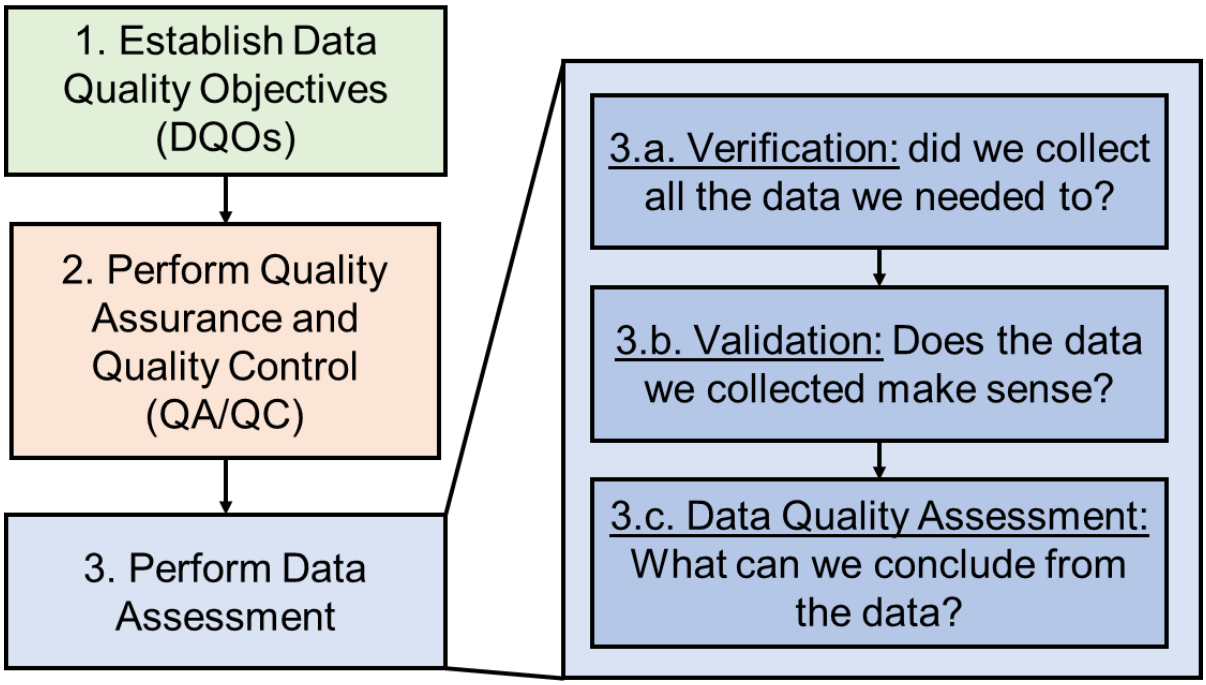


Figure 1. The data life cycle, focusing on the steps within data assessment.

2.1 Planning

The planning phase starts with the DQOs process, which asks: "What information is needed to answer the question that has been asked?" DQOs are the performance and acceptance criteria for radiological measurement data relating to a specific objective. For example, the DQOs for deciding whether to administer potassium-iodide tablets are different from those for deciding whether or not to relocate neighborhood residents. The objective may be to support a decision or to estimate a value. Each decision or estimation objective may have different types of DQOs.

- For decision objectives, the DQOs are typically defined in terms of error rates—both Type I errors (false positives) and Type II errors (false negatives)—and confidence level for a statistical test on a data population parameter of interest. An example of a decision objective is: "Determine if the average dose rate in a specific area is greater than 10 mR/hr with no greater than 5% Type I error, 5% Type II error, and 95% confidence."
- For estimation objectives, the DQOs are often defined in terms of acceptable estimation uncertainty for a certain level of confidence. An example of an estimation objective is "Estimate the dose rate to within ±0.1 mR/hr with 95% confidence."
- DQOs for both types of objectives may also include qualitative information, and in the case of decision objectives, may not have a quantitative outcome.
- Both types of objective (decision and estimation) may appear in a nuclear or radiological response scenario (EPA 2006b).

DQOs are then considered in conjunction with data quality indicators to derive the measurement quality objectives (MQOs), which are the set of measurement performance criteria that must be met for each data quality indicator. Data quality indicators can be broadly summed up as precision, accuracy, representativeness, completeness, comparability, and sensitivity, often referred to as the PARCCS parameters. MQOs are then chosen for each of the PARCCS parameters required of the measurement processes to achieve the DQOs. MQOs (Measurement Quality Objectives) can be quantitative or qualitative, depending on the DQOs they support. MQOs for radiological measurements will often include which radiation types must be measured (completeness), a detection limit value for each isotope to be met (sensitivity), and the allowable uncertainty on individual measurements (precision) (Multi-Agency 2004).

The PARCCS Parameters

Precision – "How much variation is there between measurements?" This parameter describes the reliability by which an investigator can reproduce the sample results. It measures the amount of dispersion among series of measurements and is often provided as a standard deviation.

Accuracy – "How close is each measurement to the true value?" This parameter describes the comparison of a result to a consensus value, generally expressed in terms of an error, either as an absolute value or percentage, where the measurement is compared to a mean or known true value (the latter is usually for laboratory sample analysis).

Representativeness – "Do the samples come from the same area?" This data parameter describes how well a sample characterizes or describes a specific population and is generally reported on a qualitative basis. This can be evaluated, for example, by checking whether measurements were taken on the same kind of terrain, or within a certain distance from each other.

Completeness – "Did we collect all the data we planned for?" This parameter describes how many measurements were collected compared to the number of planned measurements, generally expressed as a percentage of the actual to the planned numbers of measurements. It can also describe the completeness of the information collected for each individual measurement.

Comparability – "Can two measurements or sets of measurements be compared?" This parameter describes whether measurements can be compared based on the instrument used, the type of measurement, the units used, and the question being answered. This is generally a qualitative parameter.

Sensitivity – "Are the instruments used sensitive enough to detect what we are looking for?" This parameter describes at what threshold value an instrument can detect radiation. Instruments may have several different sensitivity thresholds for different radiation types and for different specific radioisotopes.

2.2 Implementation

The implementation phase includes the QA/QC processes. QA/QC describes the management system and technical activities, respectively, used during data collection to collect measurements that meet the MQOs. QA encompasses the data management system used to

capture and record the data life cycle, such as documentation and data collection. QC encompasses the activities performed in a QA program that measure and record applicable data quality indicators, such as instrument calibration certification and sample splitting for analysis by multiple laboratories. QA/QC broadly emphasize the PARCCS parameters in data collection, as well as an awareness of instrument detection limits and measurement range, and the practice of taking extra samples, called QC samples. Effective QA/QC ensures that MQOs are met (EPA 1996).

2.3 Assessment

The assessment phase of the data life cycle includes three main steps: verification, validation, and DQA. Verification and validation are often grouped together as "V&V" by convention, but they are separate processes and should be treated separately in data assessment guidance. In this section, each step of the assessment phase is described. These steps may be undertaken by one or more data assessors. A data assessor is an individual who performs one or more of the data assessment steps.

2.3.1 Verification

Data verification asks the question, "Did all of the requested information get collected?" The purpose of data verification is to make sure that the records associated with a specific set of radiological measurements reflect all of the processes and procedures used to generate that dataset. In ideal circumstances, a list of all data types and fields that were developed as part of the planning phase and requested during the implementation phase is provided and can be compared to the information that was collected. Such documentation is part of regular operating procedure at analytical laboratories that receive samples from the field. These laboratories typically use well-understood chemical separation methods to separate particular isotopes of interest from a larger sample and then quantify the concentration based on high-performance radiation detection instrumentation. The records of sample collection and analysis, chain-of-custody, and instrument calibration records must all be maintained as part of the incident file to show that the measurement was performed as intended in case decisions based on these measurements are challenged (EPA 2002, ANSI 2012).

Field measurements may also have a specific set of requested information fields that measurements can be compared to. For example, measurements for the objective of determining whether radioactive contamination is present will require multiple measurements including the measurement value and units, the location, the background radiation level at the location of interest, and the time and day of the measurement, among other things. If this basic information is not available, then a data assessor might consider the measurement data of insufficient quality to address the objective.

2.3.2 Validation

Data validation asks the question, "Were the MQOs achieved for each measurement and as a whole?" The purpose of data validation is to compare measurements to the PARCCS parameters and determine if the requirements for those parameters have been achieved by the individual measurements being considered. Much like verification, in ideal circumstances, measurement results can be compared to a list of all MQOs established during the planning and implementation phases. For example, when trying to measure the radionuclide concentration in an area, a measurement from a personal radiation detector (PRD) may not have the precision required to estimate the concentration of different radionuclides with sufficient uncertainty. As

another example, an instrument's sensitivity range setting may have been set too high to accurately detect the quantity of radiological material of concern. In each case, a data assessor might then consider the measured values invalid for the objective being addressed.

If an individual measurement does not meet one of the MQOs, or if other observations about the available information are made, a signifier or flag is typically applied to the measurement to indicate what the particular issue or observation about the available information is (EPA 2002, ANSI 2012). This means that some type of label has been applied to the measurement to facilitate review against its intended use. This type of labeling also helps in the review process when multiple individuals are involved.

When flagged data are transferred from the person performing validation to quality assessment or other analysis, the person receiving the data will be able to review the flags on each data point to determine if the quality of the available information regarding the collection of that measurement value is suitable for their needs. The specific nomenclature of the flags that are applied to data may vary between organizations. For example, the American National Standards Institute and the American Nuclear Society have produced a standard for "verification and validation of radiological data for use in waste management and environmental remediation" (ANSI 2012) in which an (R) flag indicates that a measurement is rejected or unusable, and a (J) flag indicates that a measurement is estimated. In other cases, case narratives will be included with measurement results that provide validation information in an explanatory form. Flags and case narratives are sometimes generated for field measurements but are much more commonly applied to results from analytical laboratories.

2.3.3 Data Quality Assessment

DQA asks the question: "Can the available information be used to address the objective, and what conclusions can be drawn?" The purpose of a quality assessment is to determine whether the measurements collected can be used to address the objective based on the results of the V&V steps. If so, measurements can be used to draw conclusions based on the available information (EPA 2000). The entire assessment phase of the data life cycle may sometimes be referred to as DQA, but this is not technically correct because it omits the distinct V&V steps.

The DQA step of the assessment phase may involve the use of statistical tests to draw specific conclusions about a set of measurements with a certain percent confidence, and false positive and false negative error rates, given a certain amount of standard deviation in the measurement set. When applicable and practical, such statistical testing can be a powerful and defensible method for drawing conclusions about a set of measurements (EPA 2000). However, using such methods also requires careful translation of statistical concepts and meaning to response officials who may not have a technical background. In addition, depending on the confidence and error rates specified, such methods may require a large number of measurements or samples, which may be cost-, resource-, and time-prohibitive for some applications. Finally, sometimes statistical methods simply may not support the objective at hand, such as when complete scanning of an object is possible. In such cases, DQA may be reduced in scope to a review of the outputs of the V&V processes and a flow chart or check list for the data assessor to help draw conclusions.

This is the final step of the overall data life cycle, but as discussed above, it may result in a return to earlier assessment phase steps, or even a return to earlier data life cycle phases (EPA 2000). For example, a data assessor may be trying to determine whether an area is contaminated with radioactive material or not, by comparing measurement values to an action

level of concern. For this example, it is decided that the action level is three times the background radiation level. Upon completing the V&V steps, the data assessor reviews the radiological measurement data and determines that the background radiation levels in the area are not well known, but proceeds with the assessment anyway, assuming an average value. Upon completion of the analysis, the data assessor discovers that many measurement values are nearly three times the estimated background, and some are greater. An individual data assessor may draw different conclusions, as follows.

- The data assessor may conclude that because some values are greater than three times background, the area should be considered contaminated.
- The data assessor may conclude that, because the background is not known and was estimated based on average background levels from other locations, a better estimate of the background radiation level for the area should be obtained before drawing a conclusion about whether the area should be considered contaminated.
- The data assessor may conclude that some areas are contaminated, and some are not, and that more measurements should be conducted in the area to help distinguish the part of the area that is contaminated from the part that is not.
- The data assessor might combine all three conclusions: the area should be considered contaminated until a better background estimate can be obtained, and more measurements can be taken. Whether or not these actions are taken is typically left to incident command, who are usually different individuals than the data assessor.

3.0 Consequence Management Incident Response

Data quality is regularly applied in the context of large environmental remediation and decontamination and decommissioning projects. Depending on scale, these projects are typically conducted over the course of years or even decades and may involve many hundreds of people performing data quality tasks alone. For these projects, it must be thoroughly demonstrated that the areas addressed are below regulatory thresholds and can thus be released for use. Decisions about the state of an area are made only after much analysis and assessment of the measurement data.

Incident response presents an entirely different set of challenges to data quality compared to more typical applications. Incident response is a fast-paced, rapidly evolving environment where decisions must be made quickly and supported with conclusions about the available data to avoid adverse effects among the population affected by an incident. Incident impacts could be immediate or delayed and could require data to be preserved long after the early and intermediate phases are over to adjudicate injury claims and be contributed to dose assessment studies (DHS 2016, FRMAC 2010, DHS 2017). However, data quality is still important in this environment since collecting information of sufficient quality is a critical step to making well-informed, defensible, and confident decisions. For example, measurement data with high statistical variance or unknown collection parameters are less reliable than those with low variance and known collection information. Examples of collection information include the height above the ground, the date and time, background values, and instrument settings. Such information is also referred to as contextual information. It is important to consider, then, how to adapt the data life cycle to the fast-paced needs of incident response. To maximize the impact of responders' efforts, their measurements should provide the highest possible quality data without adversely affecting the effectiveness of the response or compromising the safety of responders. To that end, the process of producing high-quality data should be as streamlined, clear, and easy as possible for responders before they ever arrive in the field.

The data life cycle is primarily affected by incident response in three ways:

- the response phase,
- the technology used, and
- interoperability between response teams.

The most significant impact to the data life cycle during incident response is the response phase. The limitations on time and resources during the early phase are the most significant. V&V is a resource-intensive process, and it is unlikely that the resources necessary for a rigorous V&V process will be available in the time-critical early phase. It is likely that the initial radiological measurements will be collected using PRDs, some of which may display an exposure rate, and others may only display a single number and provide a visual, auditory, or vibratory alarm if the exposure rate exceeds a preset threshold. Yet, a decision must still be made and should be supported by available information. As the response continues, more and greater numbers of sophisticated and sensitive instruments will be brought to bear, such as ionization chambers and high-resolution gamma spectrometers, which are deployed to the incident, measure samples collected at the incident, and measure samples that are shipped to analytical laboratories. These measurements require a more thorough review because more stringent requirements (MQOs) are applied, compared to measurements collected to support

decisions in the early phase. Therefore, a modified data assessment process should be implemented for the early and intermediate phases of a response for elected and appointed officials to be as confident as possible that the conclusions drawn about the measured radiological data are accurate. Preestablished data quality requirements (i.e., DQOs and MQOs) for a specific response objective may not be available in a particular incident response scenario, as each scenario is unique in terms of the release and objectives to be addressed. However, there are some objectives common to nearly every response (e.g., calculating minimum detectable activity or count rate needed to meet a release threshold under certain conditions) for which data quality requirements can be at least partially determined beforehand. These existing requirements for commonly asked questions can be used as a starting point and modified to fit the particular scenario.

Technology also plays a role in the quality of measurement data that are collected to support specific response objectives. As mentioned above, information about the context and conditions of each measurement is required for data assessment. For example, it is important for the location of each measurement to be recorded, ideally in geospatial coordinates, so that data assessors can decide whether the measurement is useful for addressing objectives about specific areas. It is also important to collect contextual information about the conditions in which the radiological measurement data were collected, such as the height above the ground, the date and time, background values, and instrument settings. Some instruments may only provide a measurement value, and the individual recording the measurement must be cognizant of all the other required information that needs to be recorded. More sophisticated instruments may automatically record settings, date and time, and location information for each measurement and upload them directly to a database of measurements. The data quality of the measurements collected by different instruments may be the same for a given objective, but when instruments do not automatically record settings or upload them to a database, it is left to the skill and memory of the operator to select the correct settings and record the information, which leaves much room for recording errors. More sophisticated instruments may also encounter errors, such as corrupted or interrupted transmissions that also affect data quality, or issues in autocalibration and scaling, but in general the number of ways in which the measurement information can be misrecorded are fewer. However, a survey form and survey map should accompany electronic measurement data to document any variations that are made in the field that would not be recorded by an automated system.

Finally, the interoperability between field and off-site assessment and reachback teams is critical to producing quality radiological measurement data quickly. Field teams will collect measurements and transmit them to off-site assessment and reachback teams that analyze the measurement data remotely at emergency operations centers (EOCs) and national laboratories. Both sets of teams are responsible for different parts of the data life cycle. Field teams will primarily be responsible for QA/QC processes and methods, where information about instrument settings, calibrations, and the conditions in which the measurements were obtained are recorded. The off-site assessment and reachback teams are responsible for the data assessment process, as well as other analysis tasks, that use the measurement data collected by the field teams. Since the off-site assessment and reachback teams are not able to observe the measurements as they are recorded, they are entirely reliant on the reported values and contextual information provided by the field teams. Therefore, field, off-site assessment, and reachback teams should coordinate to the fullest extent possible, before and during an incident, to increase the likelihood that all data requirements are met.

In the early and intermediate phases of a response, FSLTT responders will collect a significant number of radiological measurements, including gross count and exposure rates measured from

handheld instruments, and will continue to do so even after federal assistance is available. Later in the response, high-resolution gamma spectrometers will be used for field survey measurements and for laboratory analysis measurements. Field and off-site assessment and reachback teams will need to coordinate to use the right instruments to provide the right measurement data for the right objective. The higher the quality of the measurement data relative to a specific response objective, the more useful it will be for other FSLTT agencies and the greater utility it may have in addressing other response objectives.

This page intentionally left blank.

4.0 Data Verification and Validation Methods and Processes

The V&V process does not concentrate on decisions, but on specific sampling and analysis processes and results. V&V involves investigating whether or not the project-specific MQOs for precision, bias, or other PARCCS parameters have been achieved. Whenever new measurement data are considered for supporting a response objective, personnel should begin by entering the data in the V&V process. New data include both recent measurement data that have been collected specifically to address the objective in question (e.g., “Has this area or item been adequately decontaminated?”), and existing measurement data that were collected to support a different objective but now are being applied to the current objective in question. Because each area, object, and situation is unique, each objective is unique, even if only in one way. Thus, all data that are used to support that objective must be subjected to the data assessment process specific to that objective, starting with V&V. Existing V&V guidance for verification and validation of radiological laboratory samples can be found in

- *Guidance on Environmental Data Verification and Data Validation,* EPA QA/G-8 (EPA 2002)
- *Verification and Validation of Radiological Data for Use in Waste Management and Environmental Remediation*. ANSI/ANS-41.5-2012 (ANSI 2012)
- *FRMAC Operations Manual* (FRMAC 2010)
- *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)* (Multi-Agency 2000)
- *Multi-Agency Radiological Laboratory Analytical Protocols Manual (MARLAP): Assessment, Planning, Implementation* (Multi-Agency 2004)

For both verification and validation, the adjusted process for incident response will not be as thorough and should focus on only what is absolutely required to support an objective during incident response, instead of requiring all of the process and contextual information typically required in published guidance. Details for how each process should be conducted during incident response are discussed in the follow sections.

4.1 Verification

Data verification is ideally performed whenever measured values or datasets are transferred from one individual or organization to another. The verification process generally involves checking that the measurement data collected meet the requirements for completeness, correctness, consistency, and compliance for the study or objective being addressed. The output of this process is a list of all completeness, correctness, consistency, and compliance requirements and whether or not they were met by the measurement data provided. An illustration of this process is provided in Figure 2.

> “**Data verification** is the process of evaluating the completeness, correctness, and conformance/compliance of a specific data set against the method, procedural, or contractual requirements.” (EPA 2002)

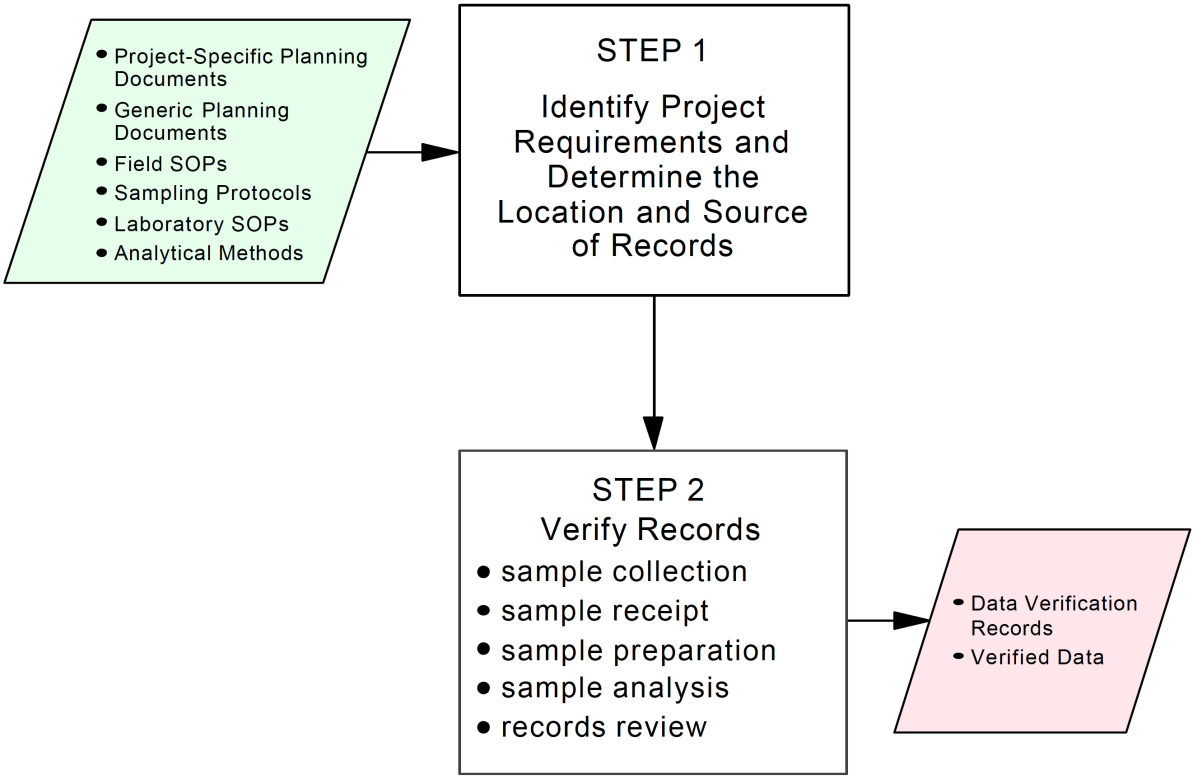


Figure 2. The data verification process steps, inputs, and outputs (EPA 2002).

The description of the verification process also assumes that data quality requirements have been developed before measurements are collected. This process is common for both sample analysis at an analytical laboratory and data collection in the field, although less guidance is available for the verification of field measurements. In such cases, a specification package is provided to the laboratory that includes the MQOs, analytical method, and other documentation. To perform verification, the specification package is compared to the data package provided by the analytical laboratory after performing the analysis, and each requirement is identified as having been met or not met. If a requirement is not met, the laboratory may be required to re-analyze the sample, or it may be able to provide or correct the information in its data package. Appendix C provides an example of poorly completed documentation.

Where a given requirement is not met, two different courses of action may be available, depending on the type of unfulfilled requirement. If the requirement calls for information that is missing or that was not provided (e.g., the units of the measurement or the uncertainty in the measurement value), the data assessor can simply ask for the missing information. If the information exists, it can be sent, and the requirement can be met. But if that information does not exist, or if it is flawed in some manner, the data requirement might remain unmet and will be noted as such.

For example, a data assessor finds that the uncertainty values were not provided with the measurement data and marks the requirement “not met.” The data assessor contacts the laboratory and discovers that the uncertainty analysis was not performed. The laboratory runs the analysis and provides the uncertainty, which then meets the requirements. The data assessor then marks the requirement “met.” If, instead, the wrong analytical method was

performed, the laboratory might need to restart its entire process, or new measurements might need to be collected if the incorrect process consumed the samples.

In many ways the requirements for field survey data will be similar to those for laboratory analysis data. However, trained technicians conduct laboratory analysis measurements in highly controlled, documented, and repeatable conditions, using established procedures. Field surveys, on the other hand, are often performed when the opportunity presents itself, often under unpredictable and changing conditions, by personnel who might lack formal training and experience. Thus, measurement data from field surveys require contextual information on the conditions in which they were collected to meet requirements for even the least stringent response objectives. For example, the range setting on a dose meter measurement should be known to have confidence that the measured value is accurate. In the case that the range was set higher than normal and the person collecting the measurement did not read their instrument correctly, the reported value might be lower than the real value by a factor of 10 (or more), which has serious consequences for responder and public safety.

The adaptation of the verification process to radiological incident response requires that

- verification requirements be developed for measurement data collected from field survey instruments, and
- a simple method for quickly verifying the completeness, correctness, consistency, and compliance of the provided information be developed.

While verification for nonemergency uses is extensive and thorough, a simpler and faster verification process should be implemented in the early and intermediate phases of a response, where thorough requirements may not exist, and where decisions will need to be made—even when detailed information is sparse. The flagging method typically applied in the validation step can be used for this purpose (see Section 4.2). For verification in this scheme, data assessors check whether or not the supplemental contextual or required information has been provided and may apply two different flags to indicate missing information based on the specific response objective: “Suspicious” or “Incomplete”, respectively. The “Suspicious” flag indicates that a data assessor might have a reasonable objection to using the measurement, but the missing information is not strictly necessary for the use of the data point to support the objective. The “Incomplete” flag indicates cases where the presence or absence of information may significantly influence a protective action decision or worker safety. The application of these flags may change, based on the response phase and the specific response objective. For example, it might not be vitally important to know the name of the person performing a survey or the agency they report to. On the other hand, a survey result of “574” cannot be evaluated or used at all unless the units are reported because of the possible difference in significance of the value, as follows:

- 574 µR/hr is elevated, but calls for few radiological safety measures,
- 574 mR/hr requires a degree of care to minimize exposure,
- 574 R/hr is dangerous and should not be entered, and
- 574 cpm might call for decontamination.

An example is provided in Table 1, where the anticipated data collected are listed on the left and the recommended flag that be applied if that information is missing is listed on the right. This example is based on data collected to address Tactic 2 in the *Radiological Dispersal Device*

(RDD) Response Guidance: First 100 Minutes (DHS 2017). An example using this table is provided in Appendix B – Data Assessment Process Example.

Table 1. Example of possible verification flags recommended for evaluating reports of contamination in the early phase.

Data	Flag, if data are missing
Name of the individual	Suspicious
Organization of the individual (if applicable)	Suspicious
Location	Incomplete
Orientation of instrument to nearby surfaces	Suspicious
Distance from instrument to nearby surfaces	Suspicious
Type of material on nearby surfaces relative to the detector	none
Time of day and date the measurement was taken	Suspicious
Instrument make and model, and of associated probes	Incomplete
Instrument serial number	none
Instrument calibration or certification records	none
Conversion factor from counts per minute (cpm) to disintegrations per minute (dpm) for the instrument (if applicable)	none
Instrument settings	none
Survey method – stationary or scanning (an instrument in a moving vehicle is considered a scan)	Suspicious
If a moving survey method was used, the approximate speed of the meter or vehicle the meter was in	Suspicious
Background measurement	Suspicious
Measurement value	Incomplete
Measurement unit (if applicable – some instruments give numbers which correspond to dose rate thresholds – see Using Preventative Radiological Nuclear Detection Equipment for Consequence Management Missions (Buddemeier et. al 2017)	Incomplete if dose reading; none if from a BNC NucAlert or D-Tect MiniRad-D
Number of measurements taken	Incomplete
Alarm type (if applicable)	none
Isotopes identified (if applicable)	none
Isotope identification confidence (if applicable)	none
Shipping manifest (if applicable)	none

4.2 Validation

Data validation is performed after data have been verified. The data validation process generally involves evaluating measurements individually and as a whole to determine if they meet the DQOs and MQOs established. The output of this process is a report on the validity of the measured data, especially including any concerns or anomalies discovered. This process is illustrated in Figure 3.

"The goals of **data validation** are to evaluate whether the data quality goals established during the planning phase have been achieved, to ensure that all project requirements are met, to determine the impact on data quality of those that were not met, and to document the results of the data validation." (EPA 2002)

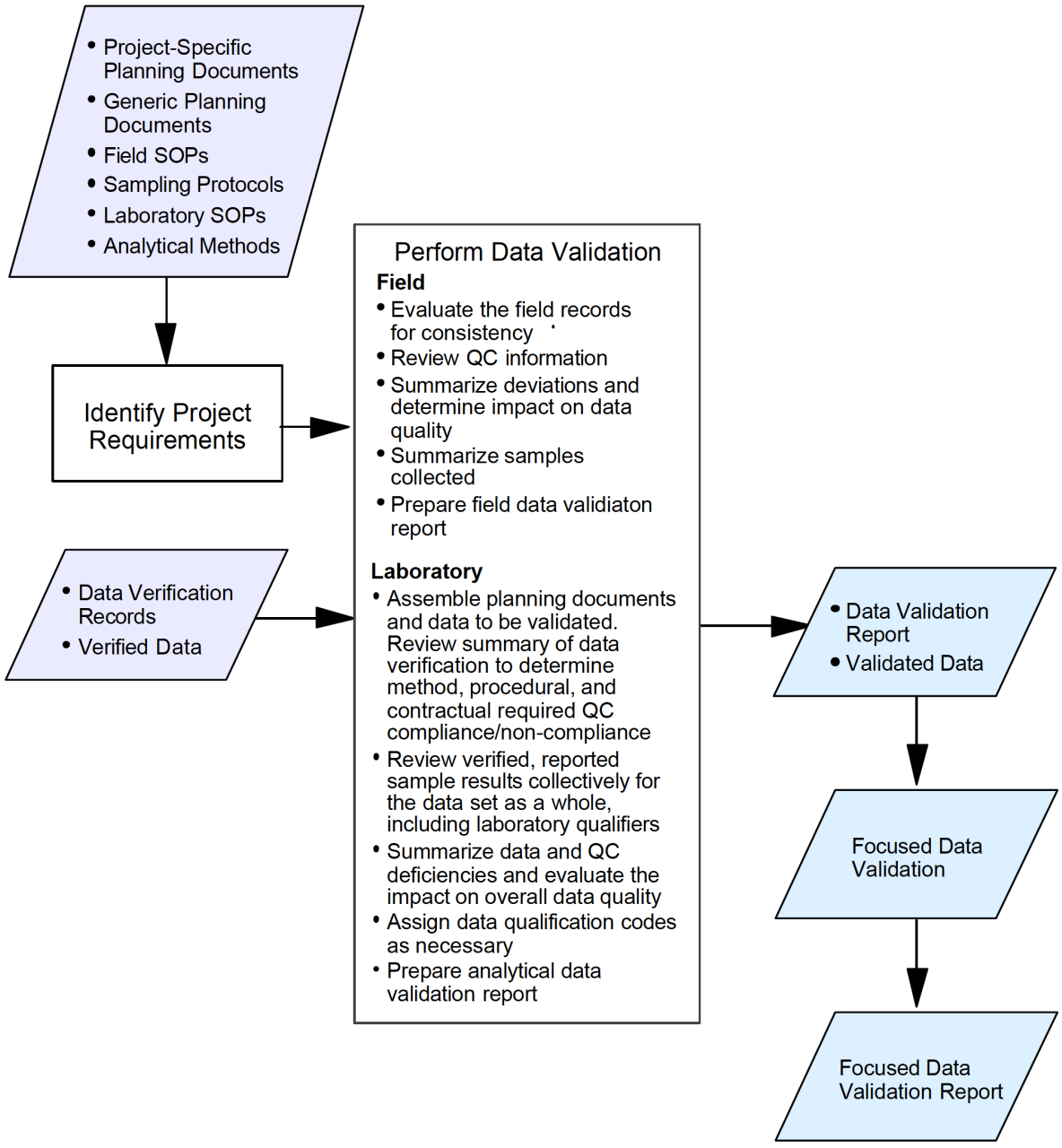


Figure 3. The data validation process steps, inputs, and outputs (EPA 2002).

As in verification, the authors describe a validation process that assumes data quality requirements have been developed as is common for sample analysis at an analytical laboratory, for example. In such cases, the data requirements will list the conditions that the

measurement data must meet to be valid. To perform the validation, the requirements are compared to each measurement individually to determine if each condition is met or not. If a condition is not met or there is some other important context that should be noted, the data assessor may flag the measurement to signify what the particular issue or observation about the measurement is (EPA 2002). The application of flags facilitates review of the data against their intended use. The specific nomenclature of the flags that are applied to the measurements may vary between organizations.

Unlike verification, the validation process should stay largely the same for incident response. Like a typical, nonemergency validation process, the application of flags notes where measurements do not meet PARCCS requirements. The use of flags also helps in the review process when multiple individuals are involved. When the person performing validation transfers flagged data to the person performing DQA or other analysis, the person receiving the validated data will be able to review the flags on the data to determine if the measurement is suitable for their needs. Unlike verification, validation flags will be more varied to describe specific issues with or observations of the measured data. In keeping with the purpose of checking against MQOs, the validation condition checks should be organized by PARCCS parameters for clarity.

Table 2 contains an example set of validation flags, where the related PARCCS parameter is listed in the first column, the action for the data assessor to complete is listed in the second column, the condition to check for in the measurement data is listed in the third column, and the flag to apply if the condition is true is in the final column. This example is based on data collected to address Tactic 2 in the Radiological Dispersal Device (RDD) Response Guidance: First 100 Minutes Guidance (DHS 2017). Appendix B provides an example using this table.

Table 2 presents flags that are suggested, based on anticipated data requirements for addressing the objective. The data assessor may apply additional flags of their own devising to account for other issues, anomalies, or observations about the measured data. The data assessor may also write the basis for their decision about a dataset, also called a case narrative, as a way of explaining the observations they made and what they may mean for the quality of the measure data.

Table 2. Example of possible validation flags recommended for evaluating reports of contamination in the early phase.

PARCCS Parameter	Action	Conditions	Flag if condition is true
Precision	Review the instrument information and measurement value and units.	The instrument does not provide a dose readout, but value representative of a dose range instead.	Imprecise
Accuracy	Review instrument calibration records, if available.	1. The calibration expiration date is earlier than the date the measurement was taken. 2. No calibration information is available.	If 1: Out of Calibration If 2: none
Representativeness	Review measurement spacing.	All measurements are within 15 meters (50 feet) of one another.	Incomplete
Completeness	Review the number of “Suspicious” flags applied to the measurement.	Five or more ”Suspicious” flags have been applied.	Suspicious Context
Comparability	Compare each measurement reading to the typical background level measured on that instrument, converting from integrated counts to count rate and accounting for the speed of the survey, if necessary.	A given measurement is less than three times the typical background for the area with the type of instrument used.	Background
Comparability	Review the location of the measurements and document anything in the immediate environment that might be responsible for increased levels of naturally occurring radioactive materials (NORM), including concrete, granite, or brick. If such materials are present, calculate how much higher above background the measurement is, and consider what isotopes were detected.	The measurement is consistent with an elevated background, and any identified isotopes are consistent with NORM.	NORM
Sensitivity	Compare the instrument settings (for example the scale, range, and so on) to the measured value.	The measured value is in the lower 20% of the full-scale range.	Invalid

This page intentionally left blank.

5.0 Data Quality Assessment Methods and Processes

Data quality assessment is the scientific and statistical evaluation of environmental data to determine if they meet the planning objectives of the project, and thus are of the right type, quality, and quantity to support their intended use. (EPA 2000)

The DQA step focuses on drawing conclusions, and whether the collected and derived data can be used to answer the question being asked reliably and defensibly. The DQA step is performed following the application of flags to the measured data in the verification and validation process. The DQA step is the process of evaluating the flags applied to measurements during previous steps to assess whether it can be used to support a specific objective or question at hand. The output of this process is either a conclusion about the data that helps answer a question, or an estimated value with known accuracy and uncertainty.

The DQA process generally follows five steps, as illustrated in Figure 4 (EPA 2000). While the authors do not provide an example of this process in full, Appendix B provides an example of the adjusted process for radiological response. The adjusted DQA process for incident response should focus on only what is absolutely required to support an objective during incident response, instead of requiring all of the processes and information typically required in published guidance.

In step 1, the data assessor reviews the key outputs from the planning phase of the data life cycle such as the DQOs, MQOs, study question or objective, and any related documents. The output from the first step is a hypothesis statement that can be accepted or rejected, based on the measurement data (e.g., "The average dose rate for this area is below 10 mR/hr"). Establishing a hypothesis is covered later in this section.

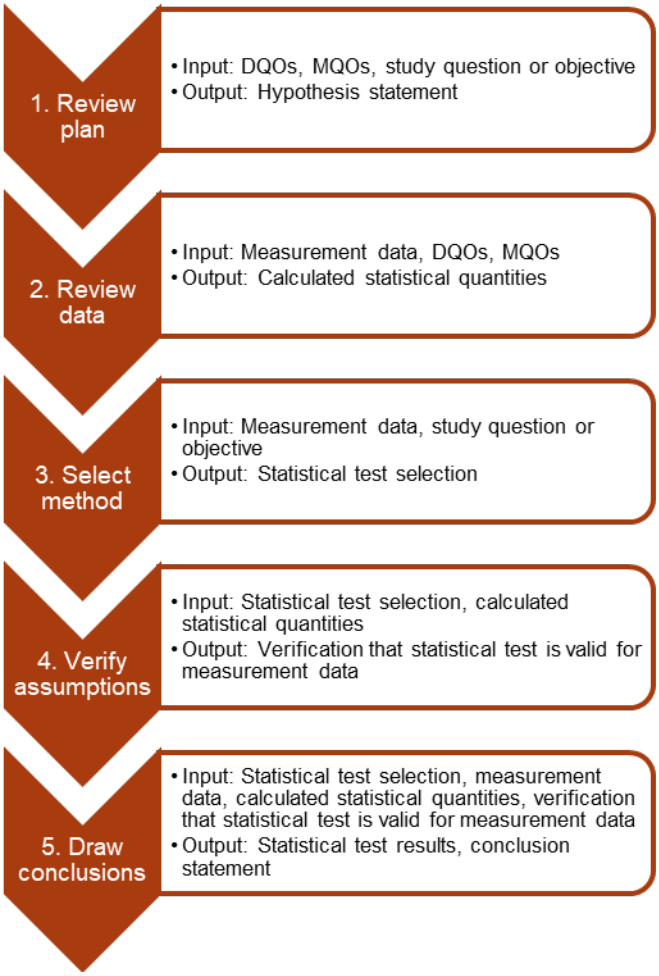


Figure 4. The five steps of the quality assessment process.

In step 2, the data assessor reviews the collected information for issues, anomalies, and observations made in the V&V steps and calculates statistical quantities and graphs the measured data.

Step 3 involves selecting a statistical method, such as a t-test, for testing the hypothesis developed in step 1. The test should be carefully chosen, based on the data available and the objective. Figure 5 provides guidance for deciding what statistical test to use.

In step 4, the assumptions of the statistical test chosen in step 3 should be compared to the statistical parameters calculated in step 2 to determine whether the data meet the assumptions required to perform the statistical test. If so, then the data assessor moves to step 5, where the statistical test is performed on the measured data. Step 5 output is a conclusion about the measurements pertaining to the objective of the study.

Detailed information on how to perform the analyses of each of these steps are provided in EPA QA/G-9S, “Data Quality Assessment: Statistical Methods for Practitioners” (EPA 2006a), and a more general explanation of these tests and how to apply them are discussed in EPA QA/G-9, “Practical Methods for Data Analysis” (EPA 2000).

Statistical tests are designed to provide a binary answer to the question of whether a given hypothesis can be rejected. This is done by first establishing a null hypothesis. The null hypothesis is the condition that is tested. The “alternative hypothesis” is then defined as the inverse of the null hypothesis. These two hypotheses must be mutually exclusive. These tests are used to show there is evidence for or against the null hypothesis. The outcome of these statistical tests is either (a) there is sufficient evidence to reject the null hypothesis (support the alternative hypothesis), or (b) there is insufficient evidence to reject the null hypothesis. Failing to reject the null hypothesis does not prove or demonstrate that the null is true, only that there is not sufficient evidence that the alternative is likely given the available measurement data. This is a subtle but important distinction in how these statements are framed.

For example,, if one is trying to determine if the public can be allowed to have uncontrolled access to an area, then the null hypothesis would be that the average dose rate is *greater than or equal to* 10 mR/hr (the regulatory limit for permitting uncontrolled access to an area), and the alternative hypothesis is that the average dose rate in an area is *less than* 10 mR/hr. If the results of the statistical test indicated that the null hypothesis could not be rejected, this would not necessarily mean that the average was greater than or equal to 10 mR/hr. It could be that more measurements would be needed to achieve the percent confidence required, perhaps because the standard deviation of the measurements was greater than expected.

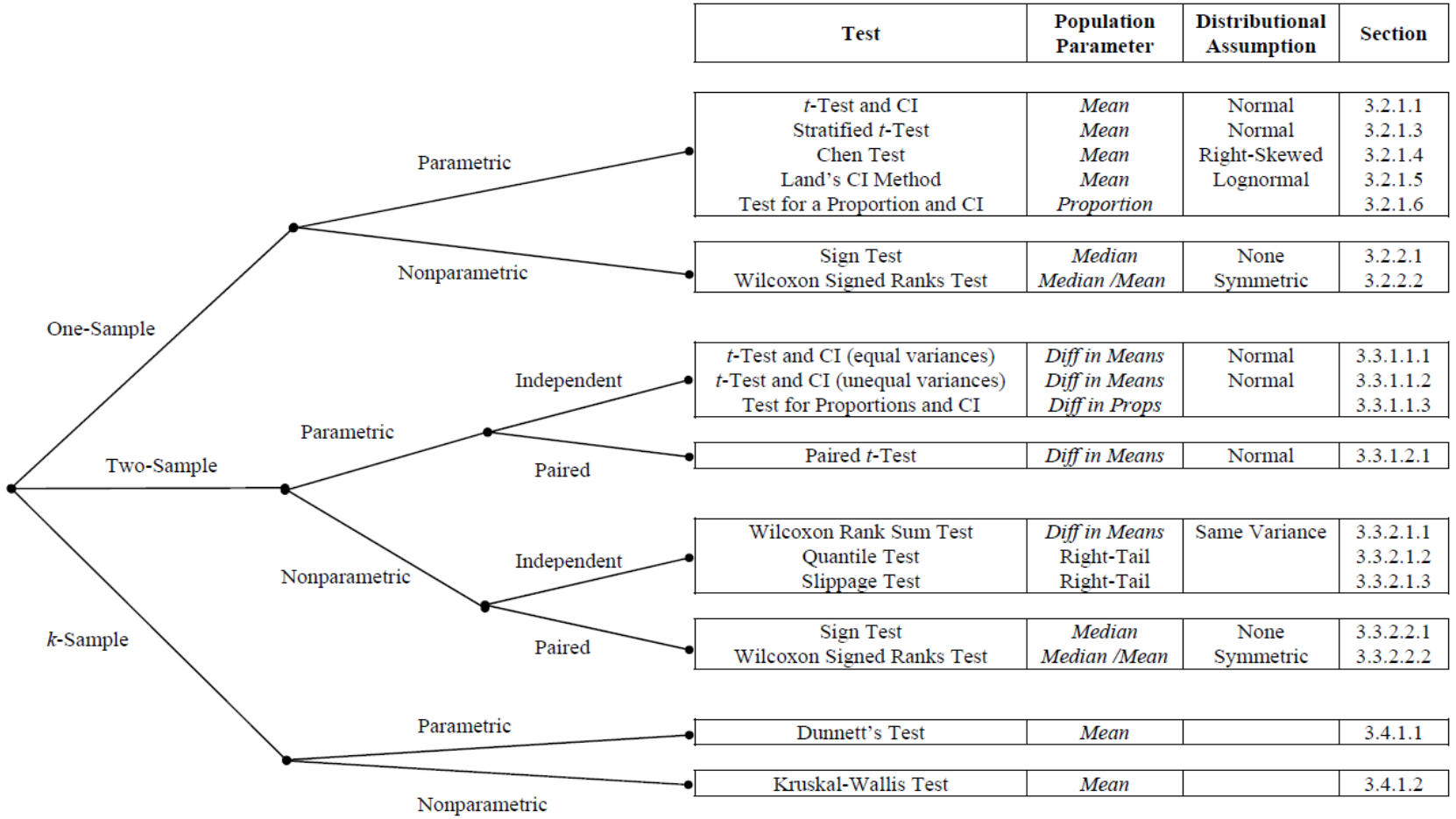
The way in which the null and alternative hypotheses are selected is important. Because of the subtlety described above, *the null hypothesis should always be set up as the more conservative case.* Consider a reverse of the setup described in the previous paragraph: the null hypothesis is that the average dose rate is *less than* 10 mR/hr , and the alternative hypothesis is that the average dose rate in the area is *greater than or equal to* 10 mR/hr. Suppose after conducting the test, the results indicate that the null hypothesis cannot be rejected. This leaves the likelihood of both hypotheses inconclusive—one is unable to reject the null hypothesis, but there is insufficient evidence to accept it.

The data assessor should also be concerned with possible false rejection error. False rejection error, or “Type I” error, is when the null hypothesis is incorrectly rejected as the result of a statistical test. False rejection error arises because the true distribution of values that could be

measured is unknown. If the distribution is not characterized well enough, false rejection error is more likely. In the first example above, suppose three measurements were collected in three discrete locations in the area of concern: 9.9 mR/hr, 9.8 mR/hr, and 9.7 mR/hr. The average of the three values is 9.8 mR/hr, which is below the 10 mR/hr dose rate limit. However, only three measurements were taken—it could be the case that many other locations in the area are above 10 mR/hr and simply were not measured. Suppose that two other measurements could be collected from the area that would yield values of 10.2 mR/hr and 10.3 mR/hr. When all five values are considered, the average dose rate is 10.1 mR/hr. In this case, the original null hypothesis, that the average dose rate is *greater than or equal to* 10 mR/hr, would be rejected based on the measurements. However, the true average dose rate is of 10.1 mR/hr means that the null hypothesis cannot be rejected.

One can control false rejection error by specifying a percent confidence for statistical testing. The percent confidence is the probability of correctly concluding that the null hypothesis should be rejected and is the inverse of the false rejection error rate. For example, if the false rejection error rate is 5%, then the percent confidence is 95%. Choosing the percent confidence for a statistical test is a matter of risk tolerance and may be unique to a given situation, data assessor, and elected or appointed official making a decision. The 95% confidence level is a widely used default. Figure 5 is a decision tree that can help in selecting the appropriate statistical test for the available data.

A wide array of statistical tests, assumptions, hypotheses, examples, and limitations are provided in the EPA QA/G-9 series of documents and will be used to design guidance for relevant questions and scenarios that can be addressed by responders in a radiological incident. The QA/G-9S document in particular describes a wide variety of statistical tests in detail, including information that helps the reader select the appropriate test to use based on assumptions that must be valid to use it, as shown in Figure 5 (EPA 2006a). These methods are the basis for the guidance in the *Multi-Agency Radiation Survey and Site Investigation Manual* (MARSSIM) (Multi-Agency 2000) and *Multi-Agency Radiological Laboratory Analytical Protocols Manual* (MARLAP) (Multi-Agency 2004), manuals published by multiple federal agencies, and therefore make a sound starting point for application to incident response.



Test	Population Parameter	Distributional Assumption	Section
t-Test and CI	*Mean*	Normal	3.2.1.1
Stratified *t*-Test	*Mean*	Normal	3.2.1.3
Chen Test	*Mean*	Right-Skewed	3.2.1.4
Land's CI Method	*Mean*	Lognormal	3.2.1.5
Test for a Proportion and CI	*Proportion*		3.2.1.6
Sign Test	*Median*	None	3.2.2.1
Wilcoxon Signed Ranks Test	*Median /Mean*	Symmetric	3.2.2.2
t-Test and CI (equal variances)	*Diff in Means*	Normal	3.3.1.1.1
t-Test and CI (unequal variances)	*Diff in Means*	Normal	3.3.1.1.2
Test for Proportions and CI	*Diff in Props*		3.3.1.1.3
Paired *t*-Test	*Diff in Means*	Normal	3.3.1.2.1
Wilcoxon Rank Sum Test	*Diff in Means*	Same Variance	3.3.2.1.1
Quantile Test	Right-Tail		3.3.2.1.2
Slippage Test	Right-Tail		3.3.2.1.3
Sign Test	*Median*	None	3.3.2.2.1
Wilcoxon Signed Ranks Test	*Median /Mean*	Symmetric	3.3.2.2.2
Dunnett's Test	*Mean*		3.4.1.1
Kruskal-Wallis Test	*Mean*		3.4.1.2

Figure 5. A decision tree for selecting the appropriate statistical test method for the available data (EPA 2006a).

5.1 Early Phase Process

As with verification and validation, the DQA process must be simplified for radiological incident response. The most important adjustment will be whether to use statistical testing. During the early phase, measurement activities will be focused on basic characterization of the incident scene and area to which contamination may have been spread for immediate public and responder protection. During the early phase decisions will need to be made quickly. Only the most basic data quality measures should be taken to avoid slowing the pace of response activities. Therefore, statistical hypothesis testing may not be appropriate for the early phase, and the DQA process may omit steps 3 and 4 (Figure 4).

Data assessment in the early phase is applied by reviewing measurement data and any accompanying reports describing the conditions in which the measurement was performed. The main task for data assessors in the early phase is to determine if measurements reported indicate potential elevated levels of radiation or expected elevated background radiation levels based on nearby land and vegetation features, building materials, and so on. The data assessor must:

1. Decide to include or reject individual measurements based on flags applied during the V&V steps,

Decide on the suitability of the measurement data as a whole,

2. Draw conclusions about the available information and the state of the incident, and
3. Present these conclusions to elected and appointed officials.

Table 3 provides an example of a proposed best practices DQA process and recommended conclusions that can be applied to assess reports of contamination during the early phase. This example is based on information collected to address Tactic 2 in the Radiological Dispersal Device (RDD) Response Guidance: First 100 Minutes Guidance (DHS 2017). An example using this table is provided in Appendix B.

Table 3. Example of possible DQA steps recommended for evaluating reports of contamination in the early phase.

Condition	Possible conclusions
The “Imprecise” flag has been applied to an individual measurement.	• If no background measurement with the same instrument is available, a background measurement should be performed away from any areas suspected of elevated background levels. • If a background measurement with the same instrument is unable to be performed, additional confirmatory measurements in near same locations should be performed. • If a background measurement with the same instrument is unable to be performed, an additional “Suspicious” flag may be applied to this measurement.
The “Out of Calibration” flag has been applied.	• An additional confirmatory measurement should be performed with a different instrument in the same location. • If an additional confirmatory measurement with the same instrument is unable to be performed, an additional “Suspicious” flag may be applied to this measurement.
The “Invalid” flag has been applied to an individual measurement.	• The individual measurement should be discarded, and not counted toward the number of required measurements. • It may be necessary to collect a new measurement to replace the invalid one.
Every measurement in the dataset has had the “Background” flag or “NORM” flag applied to it.	• It is unlikely that there is a radiation hazard is at the location. • Additional confirmatory measurements from first responders may be sought. If possible, additional measurements should include gamma ray spectra for isotope identification.
The “Incomplete” flag has been applied to an individual measurement.	• The missing information should be sought from the individual who collected the measurement. • The measurement should not be considered usable until the missing information has been obtained. • Additional confirmatory measurements in the same locations should be performed.
The “Suspicious Context” flag has been applied to a measurement.	• Any missing information should be sought from the individual who collected the measurement. • Avoid using this measurement if other nearby measurements with fewer than five “Suspicious” flags are available. • Additional confirmatory measurements in near same locations should be performed.
The number of ”Suspicious” flags for a set of measurements is equal to or greater than three times the number of measurements.	• Any missing information should be sought from the individuals who collected the measurements. • Additional confirmatory measurements in the same area should be performed.
The number of measurements that do not have “Incomplete” or “Invalid” flags is less than two.	• Additional confirmatory measurements in near same locations should be performed.
Only one instrument was used to make all valid measurements.	• Additional confirmatory measurements in near same locations should be performed.
None of the above conditions is true.	• It is likely that an elevated radiation hazard is present at the location. • Compare the radiation measurements to an action level.

5.2 Intermediate Phase Process

As the incident stabilizes and priorities shift to more detailed and long-term assessment of public health impacts, the requirements for data quality will increase as the pace slows and the decisions being made require more careful answers. Statistical hypothesis testing will therefore be appropriate for many, but not all, objectives in this response phase, and the DQA process will involve all five steps (Figure 4).

Response objectives—such as releasing areas from an administrative contamination boundary and monitoring vegetation, animal products, and water supplies to determine whether an embargo should remain in place or be rescinded—will require a relatively large number of measurements to be collected to draw statistically significant conclusions. This can be approached through a number of statistical hypothesis tests such as comparing an average measurement value to an action level or comparing an average measurement value to a reference average value (e.g., of an area considered to be representative of background). While guidance for test selection exists, as shown in Figure 5, this involves calculation of several statistical parameters and graphing the measured data to verify that these data meet the assumptions for the test. For radiological incident response, it may be useful to make the assumption that nonparametric tests should be used. Nonparametric tests do not assume that measurement data adhere to any particular statistical distribution. This way, a default test may be provided for each objective such that fewer assumptions about the measurement data need to be validated. Nonparametric upper tolerance limit (UTL) and item sampling tests are expected to be particularly applicable for many response objectives. These are explained in the next two sections.

Statistical Testing and Absolute Certainty

Statistical testing can never provide 100% certainty. Data assessors use statistical tests to infer characteristics of a distribution of real values based on a limited number of measurements. Differences always exist between the samples from a population and the whole population, even when 100% of an area or 100% of available items are surveyed.

Data assessors may need to convey this concept to response leadership, including incident command and elected and appointed officials. However, methods for conveying this concept are outside the scope of this report.

5.2.1 Upper Tolerance Limit

UTL is a statistical test to determine if a certain percentage of the area of a distribution (a percentile) is below a user-defined value. In terms of response example, this could be phrased as, "Given the *number* of measurements collected, we can be *95*% confident that *95*% *of the area* is below the action level." Comparing an action level to a percentile is advantageous versus comparing an action level to a mean. If the mean is compared to the action level, then a significant portion of the area being surveyed (up to about 50%) may still be above the action level. An illustration of this potential issue is shown in Figure 6, where the mean is below the action level, but a significant proportion of the total distribution is above the action level. On the other hand, the 95th percentile is significantly greater than the action level and would indicate that much of an area is above the action level, as well.

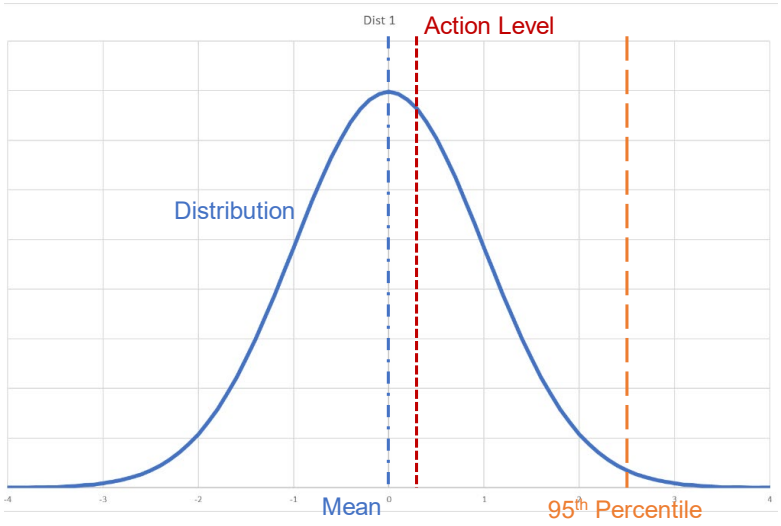


Figure 6. Example distribution illustrating the mean and the 95th percentile, and an action level they are being compared to.

The UTL test can be used when determining if an area is above or below an action level, and if it can be released for use or if further decontamination is required. The UTL test is used to evaluate whether a certain fraction of a distribution is greater or less than the action level value (Millard and Neerchal 2001). If any individual measurement is above the action level, then the null hypothesis automatically cannot be rejected. If not, then the data assessor needs to know two of the three following pieces of information:

- The fraction of the area that must be proven to be lower than the action level
- The percent confidence desired
- The number of measurements to be used for the test.

Which two pieces of information are required depends on how the test is constructed, based on the preferences of elected or appointed officials making the decision and limitations of available information. The test can be constructed to calculate three different values:

1. The number of measurements required to test whether a specific fraction of an area is below the action level with a specific percent confidence (e.g., “How many survey points are needed to have 95% confidence that we can show that at least 95% of this room can be released?”);
2. The percent confidence achieved given a number of measurements and a specific fraction of the area that is desired to be below the action level (e.g., “With the number of survey points that have been collected, how confident are we that 95% of this area can be released?”); or
3. The fraction of the area that can be concluded to be below the action level given a specific percent confidence and the number of measurements (e.g., “With the number of survey points that have been collected, how much of this area is likely to be below the release limits with 95% confidence?”).

While value 1 is useful in the planning phase of the data life cycle, values 2 and 3 will be more useful in the assessment phase, after measurements have already been collected and

significant effort must be expended to collect more. The formulation of this test is relatively simple for each of the three constructions above (PNNL 2022a):

$$n = \frac{\ln\left(1 - \frac{C}{100}\right)}{\ln(P)}, \quad \text{(Eq. 1)}$$

$$C = \left(1 - \exp\left(n \cdot \ln(P)\right)\right) \cdot 100, \quad \text{(Eq. 2)}$$

$$P = \exp\left(\frac{\ln\left(1 - \frac{C}{100}\right)}{n}\right), \quad \text{(Eq. 3)}$$

where
n is the number of measurements
C is the percent confidence level (e.g., “95” for the 95% confidence level)
P is the fraction of the area (a percentile of the distribution, e.g., 95% of an area is expressed as 0.95).

Generalizing the statement given previously, the result of this test would read: *“*Given *n* measurements collected, we can be *C* percent confident that *P*100* percent of the area is below the action level.”

It can be inferred from the equations above that, *the higher the desired percent confidence and the higher the fraction of the area, the more measurements need to have been collected*. Figure 7 and Figure 8 show the relationships between the number of measurements and the percent confidence and percent area clean, respectively. For example, testing for 95% of the area with 95% confidence requires 59 measurements, whereas testing for 99% of the area with 99% confidence requires 459 measurements. These plots may be helpful when data assessors need to explain to others the resources, time, and cost of achieving a high percent confidence about a large percentage of an area. These plots also serve as an illustration of why such statistical methods are not practical in the early phase of a response.

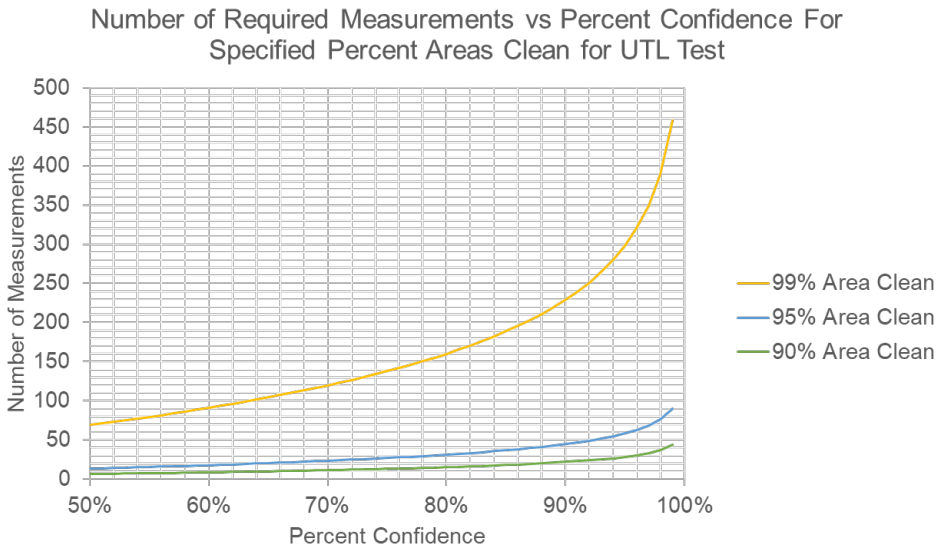


Figure 7. Relationship between the number of measurements required and specified percent confidence when testing for a given percentage of an area.

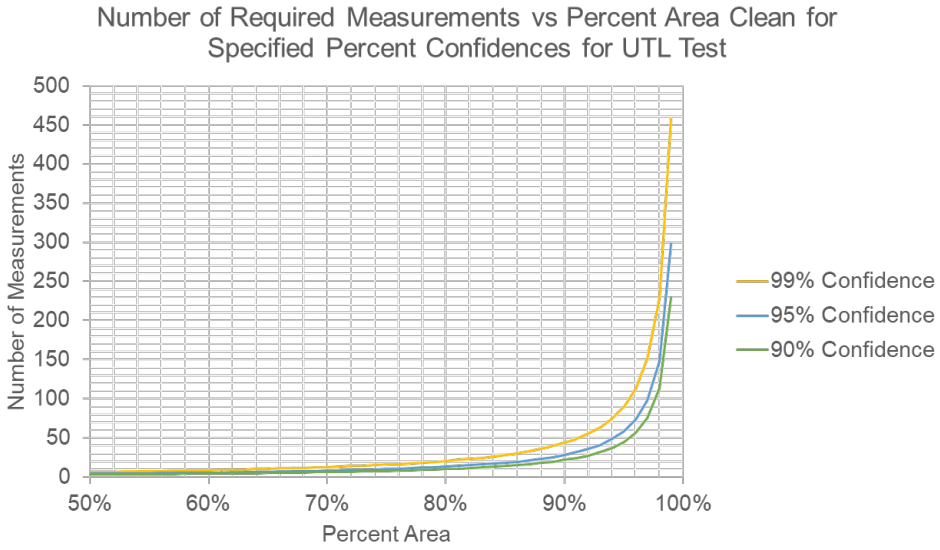


Figure 8. Relationship between the number of measurements required and percentage of an area when testing for a given percent confidence.

5.2.2 Item Sampling

The item sampling test is applicable whenever a total population of a discrete number of items is being evaluated against an action level. This test is based on an approach known as "Compliance Sampling" (Schilling and Neubauer 2009) or "Accept on zero attribute compliance sampling" (Squeglia 2008, Bowen and Bennett 1988). Similar to the UTL test, if any individual item is measured to be above the action level, then null hypothesis cannot be rejected automatically. If not, the data assessor needs to know the total number of items that represent

the total item population, how many measurements may be unacceptable (above the action level), the number of items from the total item population that must be acceptable, and either the desired percent confidence or the number of items that were measured from the total item population. Which of the last two pieces of information are required depends on how the test is constructed, based on the preferences of the elected or appointed officials making the decision and limitations of available information.

The pieces of information required to be known is reduced to three for the case that no measurements may be unacceptable, which is likely to be the case for incident response. In this case, the information that must be known is:

- the total number of items that represent the total item population,
- the number of items from the total item population that must be acceptable (often expressed as a percentage), and
- either the desired percent confidence or the number of items that were measured from the total item population.

The test can be constructed to calculate either:

1. The number of measurements required to achieve the desired percent confidence level that, given all measured items are acceptable, the number of items from the total item population are acceptable (e.g., “How many items must be determined to be below release limits in order to have 95% confidence that 95% of all items are below release limits?”); or
2. The percent confidence level that, given a number of measurements and that all measurements are acceptable, the desired number of items from the total item population are acceptable (e.g., “How certain are we that, after finding that 137 items that can be released, all 192 will be lower than release limits?”).

Like the UTL tests, item sampling test 1 is useful in the planning phase of the data life cycle, and item sampling test 2 is more useful in the assessment phase. The formulation for item sampling test 2 is (PNNL 2022b):

$$C = 100 * \left(1 - \left(1 - \frac{2n}{2N - V + 1} \right)^{V} \right), \qquad \text{(Eq. 4)}$$

where
C is the percent confidence level (e.g., “95” for the 95% confidence level),
n is the number of items that were measured,
N is the number of items in the total item population, and
V is:

$$V = \left\lceil (1 - p) \cdot N \right\rceil, \qquad \text{(Eq. 5)}$$

where p is the percentage (expressed as a fraction, e.g., “0.95” for 95%) of the total item population that must be acceptable, and where the bracket notation indicates that non-integer values should be rounded up to the next highest integer. A generalized statement of a result from this test could be given as: *“*Given n items measured of N total items, we can be C percent confident that $p*100$ percent of the total number of items are acceptable.”

The variable *V* must always be at least 1, which is relevant when 100% of the total item population must be acceptable and *p* is 1. Requiring 100% of the total item population to be acceptable will often be impractical for incident response. Similar to the UTL method, *the higher the desired percent confidence and the higher the percentage of total items desired to be below release limits, the more measurements need to have been collected*. Figure 9 and Figure 10 show the relationships between the fraction of the total number of items that must be measured and the percent confidence and percent of items needing to be acceptable, respectively. For example, *if the total item population is 1000, then a 100% total item population acceptability rate will mean that 950 items will need to be measured – nearly the entire population itself.* As with the UTL method, it is important for data assessors to communicate these mathematical relationships with response leadership.

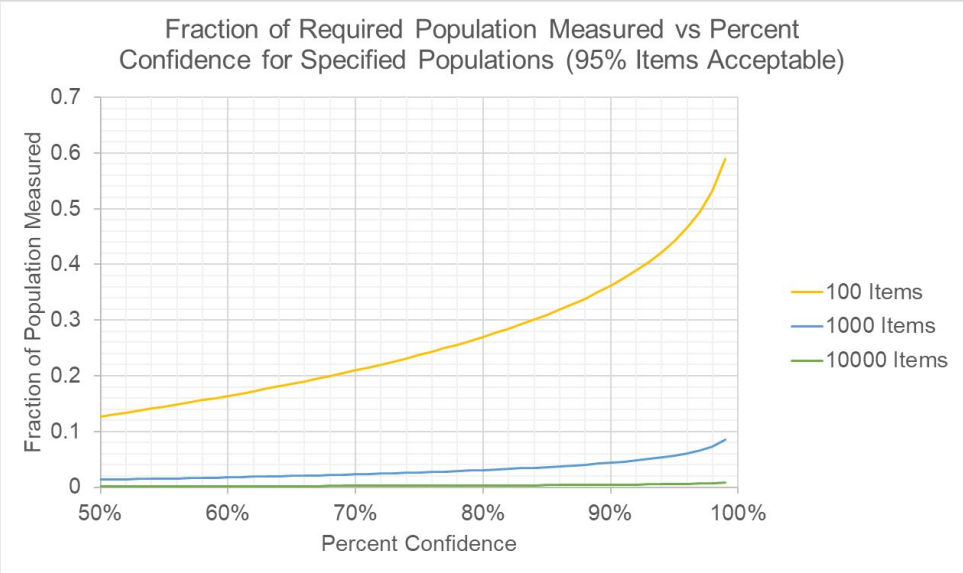


Figure 9. Relationship between the fraction of total items that must be measured and percent confidence when testing for a given number of total items.

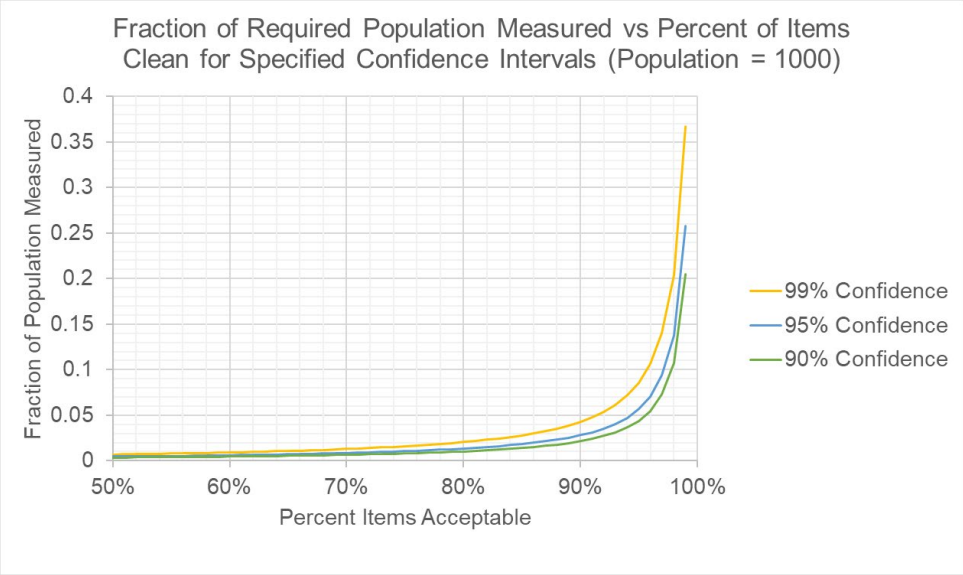


Figure 10. Relationship between the fraction of total items that must be measured and percent of total items that must be acceptable when testing for a given percent confidence.

6.0 Recommendations

The research into V&V and DQA best practices and methods informs the recommendations found in this section. The recommendations are split into two subsections: recommendations for the DQA toolkit, and other recommendations about implementation of DQA practices for radiological incident response outside the scope of this project.

6.1 DQA Toolkit Recommendations

This section covers recommendations regarding the DQA toolkit to be developed under Task 3 of this project. Each recommendation is covered by a short, boxed recommendation statement and a larger discussion and explanation following the box.

DQA Toolkit.1	An ideal DQA toolkit will include 1) a guidance document for how to apply the data assessment process to common radiological incident response objectives, and 2) an interactive electronic data assessment workbook that implements the data assessment methods described in the guidance document.

The purpose of the guidance document is to

- provide readers with an introduction to the data life cycle in general, to the assessment phase specifically, and
- describe V&V and DQA methods that data assessors can use to support emergency response officials who must make decisions that require radiological measurement data to resolve.

Since V&V and DQA procedures can be time consuming when applied to the large volumes of data generated during response to a radiological incident, FSLTT data assessors would benefit from a document that can guide them through the assessment phase of the data life cycle.

For the same reason, FSLTT data assessors would benefit from an interactive electronic data workbook to assist them in implementing the V&V and data quality steps, specifically to assist in the application of data flags and execution of the mathematical calculations for statistical tests. This toolkit workbook will help expedite the assessment phase process steps and reduce the burden on data assessors.

DQA Toolkit.2	Self-contained toolkit The toolkit guidance and workbook should not require data assessors to reference any other data quality guidance and should not require any computational tools more complex than a pocket calculator or smartphone.

The toolkit guidance document should be self-contained; that is, the document should include a full explanation of the specific assessment phase methods recommended in the document itself without the need to reference other documents. In practice, this would mean providing guidance on specific flags to apply during the V&V steps, how to interpret those flags in the DQA process,

and the equations necessary to perform any statistical tests. The reader should not need to reference other guidance to understand how to perform the tests recommended in the toolkit guidance or to understand which flags to apply. Additional reading may be required for more complex tests and any statistical calculations outside the scope of the tests recommended in the toolkit guidance itself. The guidance document should therefore also not rely on any computational tools more complicated than a pocket calculator. The most complex functions required should be limited to exponential and natural logarithm functions, which are both present in inexpensive pocket calculators and on mobile phones. Graphing calculators should not be required to perform the statistical tests in the toolkit guidance. This way, if necessary, the data assessor may perform their tasks with a minimum of other tools, and even without the toolkit workbook if necessary.

Similarly, the toolkit workbook should be self-contained, relying on no outside software or tools to apply flags to available data or perform calculations. The data assessor may need to transfer data and information from the chemical, biological, radiological, and nuclear (CBRN) Responder tool into the toolkit workbook, and the workbook should therefore be compatible with CBRNResponder data export formats. However, this transfer should not require any additional tools or software to perform. The workbook should also be able to be installed on a computer and not require any kind of connection to the Internet. These recommendations are made to make it as straightforward as possible to apply good data assessment practices, even in the emergency phase, so data assessors can perform their tasks anywhere, at any time.

DQA Toolkit.3	Organization The methods implementation of the DQA toolkit should be organized around response objectives to reinforce the idea that each objective is different and should be addressed individually.

Data quality is always relative to the decision to be supported, the question to be answered, or the objective to be addressed. Therefore, the toolkit guidance and workbook should be organized such that V&V and DQA methods are written to support specific response objectives. For each objective, the guidance document and workbook will offer default, example conclusions that may be applied to inform decisions and resolutions.

Of course, the specific objectives to be achieved in each response will be unique in one way or another and therefore will be entirely dependent on the specific response itself. Instead of trying to predict each specific response objective that may arise, the toolkit should instead focus on a set of objectives common to many incidents. The objectives that the authors recommend be addressed in the toolkit, by response phase, include:

- Early phase
  - Assess reports of elevated radiation levels
  - Compare measured radiological values to an action level
- Intermediate phase
  - Screen items
  - Release areas
  - Scan for hotspots
  - Monitor vegetation

- Monitor water
- Monitor animal products
- Estimate radionuclide concentration.

Additional response objectives may be added to the list above, based on the scoping discussion in Task 3.1 of this project. For example, closer examination of the *Radiological Dispersal Device (RDD) Response Guidance, Planning for the First 100 Minutes* (DHS 2017) may reveal what additional data will be collected and in what form and important decisions need to be made in the early phase of response. Additional discussions with state and local responders may also yield additional common response objectives that should be supported.

DQA Toolkit.4	Graded approach Statistical methods should be restricted to intermediate phase objectives and should not be used for early (emergency) phase objectives.

The DQA process is necessarily abbreviated at the start of an incident response, when responders' most precious resource is time. Only the most basic data quality measures should be taken in the early phase to avoid slowing the pace of response activities. The in-depth DQA processes outlined in the existing guidance reviewed for this research will be combined with outreach to state and local agencies that are part of the intended audience for the toolkit guidance document and workbook to inform the balance of technical rigor of the standard data assessment process with the limitations of radiological incident response.

As discussed in Section 5.0, this graded approach will be implemented by omitting steps 3 and 4 from the DQA process for early phase methods. Additionally, step 2 will be optional in both early and intermediate phases. The purpose of step 2 is to perform data exploration by calculating statistical quantities, which then inform which statistical test is most appropriate. Since no statistical tests will be performed in the early phase, step 2 is thus unnecessary for that phase. For the intermediate phase, the authors will select recommended statistical tests, such as UTL and item sampling, that support each response objective. These tests are chosen because they are nonparametric, and therefore also do not require calculation of statistical quantities, such as the mean, median, or standard deviation, because the tests assume no specific distribution. One potential downside to this approach is that, in general, nonparametric tests often require more measurements to be valid than parametric tests. Therefore, some discussion of statistical data exploration may be included in the toolkit guidance document alongside parametric tests to point data assessors to other resources for calculating statistical quantities and for describing other tests may be used to support a specific objective.

DQA Toolkit.5	Phase scope The DQA toolkit should only consider the early and intermediate phase.

While important decisions must be made and important objectives achieved throughout an incident response, the most critical and time-sensitive decisions are made in the early and intermediate phases. It is for these conditions that the typical data assessment process must be adjusted to be practical, given the time and resource constraints expected to be present during a radiological incident response. As the intermediate phase transitions to the late phase, where the incident has been stabilized and well-characterized, and recovery operations become the

main focus of ongoing efforts, the full data life cycle should be implemented. The full data life cycle for normal operations has been thoroughly described in many other guidance documents, such as those reviewed for developing this report, and has been implemented in other software tools, such as Visual Sample Plan (Matzke et al. 2014). Therefore, the full data life cycle does not need to be covered by the toolkit guidance or workbook.

DQA Toolkit.6	Audience scope The DQA toolkit should be primarily targeted toward SLTT personnel who need to analyze radiological measurement data and present conclusions about the available information to incident response leadership, such as section leaders, incident commanders, and elected and appointed officials.

While many responders will be responsible for implementing some part of the data life cycle, not all responders will be responsible for implementing the assessment phase of the process. Each step of the assessment phase (verification, validation, and DQA) may be handled by different individuals who may each also have a different position and be part of different sections in the incident command system structure (FEMA 2017). For example, verification may be performed in the field before the measured data are passed on to an EOC. In such a case, the leader of a strike team in the operations section responsible for collecting measurements may review and verify their team's information before sending it to the leader of a situation unit team in the planning section for validation and DQA.

The authors also recommend that the toolkit be targeted for SLTT personnel as the primary target audience, and federal responders as a secondary audience. There are three reasons for this recommendation:

- Some early phase methods may not apply to federal responders, who will not be at the scene immediately and may not be supporting the response even remotely as early as SLTT personnel.
- Federal responders may be required to implement more complex statistical methods and perform a more thorough data assessment. While the toolkit may serve as a useful reminder of basic statistical methods and the process overall, it is likely that the toolkit will not meet all of the data assessment needs of federal responders.
- Several parallel efforts are ongoing and supporting federal responder assessment phase needs.

Though the toolkit should be targeted for personnel who will fill the data assessor role, the Toolkit guidance may be useful for other responders, as well. For responders performing data collection, the guidance will provide an understanding of the data input fields, data types, and number of measurements required to support response objectives. For section leaders, incident commanders, elected and appointed officials, as well as liaisons and public information officers, the guidance will provide an understanding of what decision input is required for a data assessor to draw conclusions, and what the conclusions likely to be provided by a data assessor are.

As a general rule, the toolkit guidance should be written using little to no jargon, explain data quality terminology used, and use a straightforward, present-tense tone. The written format should include bulleted or numbered lists where appropriate, separate text boxes to explain

important concepts that might otherwise interrupt the main body flow. Graphics should supplement and reinforce the written concepts.

DQA Toolkit.7	Workbook format The data assessment toolkit workbook should be based on existing software of verified quality.

The toolkit workbook will assist data assessors in applying data quality flags and performing calculations for statistical tests that will help inform response leadership make decisions during a radiological emergency response incident. The outputs and results displayed in the workbook should be evaluated against a software quality assurance standard and a software quality assurance plan should be implemented to guide compliance with the standard chosen. Achieving any software quality standard will be made easier by basing the toolkit workbook on existing verified-quality software to avoid issues with the workbook program crashing, erasing data, or otherwise malfunctioning in computer application context. PNNL will still be responsible for verifying the quality of the calculations programmed into the base software if, for example, software such as Microsoft Excel is used to set up calculations and provide outputs.

Another option is to leverage features in existing data quality or statistical calculation software, such as Visual Sample Plan (Matzke et al. 2014) or Origin (OriginLab 2022). Using this option may be preferrable from a software quality standpoint as such software tools are already of a verified quality. The responsibility of PNNL in such a case would mainly shift to writing instructions for using the software in the appropriate way, but PNNL would still be responsible for verifying to some degree that the outputs produced in such software agree with hand calculations. Potential issues with this approach are that responders may be unfamiliar with tools like Visual Sample Plan and Origin, and tools like Origin (which requires a license fee) may not be universally available.

6.2 General Data Assessment Recommendations

This section covers recommendations regarding DQA implementation in radiological response. These recommendations fall outside the scope of the DQA toolkit project and are not intended as scope modifications or additions. The recommendations in this section are intended as suggestions and ideas for better implementation of DQA practices and methods spanning many aspects of radiological response. These recommendations complement the DQA toolkit. All recommendations are applicable to NUSTL but may also be applicable as combined efforts with partner organizations, including the Federal Emergency Management Agency (FEMA) and the National Nuclear Security Administration (NNSA). Each recommendation is covered by a short, boxed recommendation statement and a larger discussion and explanation following the box.

DQA Implementation.1	CBRNResponder The data assessment process should generally not be implemented as part of the CBRNResponder tool.

The CBRNResponder tool (formerly separated into RadResponder, ChemResponder, and BioResponder) is a data collection and geospatial display tool developed by a combination of federal agencies to provide a free-to-use platform for data collection and consolidation by SLTT organizations during an emergency response incident. This platform also includes a resource

library that contains guidance for planning, preparation, and operations during a CBRN incident. The popularity of this platform has grown over time, and it now includes many additional features advocated for by the response community to assist with data collection and management.

However, the nature of the data assessment process means that the implementation of the assessment phase steps in CBRNResponder should at least be limited to the verification step, and ideally be separate from CBRNResponder entirely. CBRNResponder is primarily a data repository and is not ideal for conducting data analysis, including data assessment. Data quality is specific to the objective at hand, and not the measurement dataset. Even verification flags that should be applied to measurements depend on the response phase and the specific response objective being supported. It will be more effective overall to copy the information from CBRNResponder into a separate database or download it into a local directory on a computer. There, the data assessor will apply operations to the data to perform their V&V and DQA tasks, which will be different for each specific objective for each specific response. Because the changes and operations will be applied in a copied dataset, it will be less likely to be misinterpreted or misused by others.

DQA Implementation.2	Training and drills Develop and deliver training and proficiency drills to SLTT first responders and data assessors on the techniques and scenarios to enhance their skills in V&V and DQA.

As with any new initiative, training incumbent and new SLTT responders and response leadership will deliver the right knowledge, skills, and attitudes to prepare them to function during the fast-paced, ever-evolving incidents without encountering delays while learning on the job. Beyond classroom instruction, a set of proficiency drills should be developed to continue data assessor development and qualification. An example list of training topics is shown in Figure 11.

The February 2020 Data Assessment Drill using RadResponder was informative for both FRMAC and SLTT responders (Fournier et al. 2020). For example, this exercise illustrated to some organizations that they needed to implement a more rigorous data life cycle, which would improve data quality overall. Such drills would also provide an opportunity for SLTT responders to practice using the DQA toolkit. Federal response agencies would need to fund and host drills, with participation from SLTT and possibly with other partners. CBRNResponder would need to be featured in all drills since this is where SLTT responders will upload data.

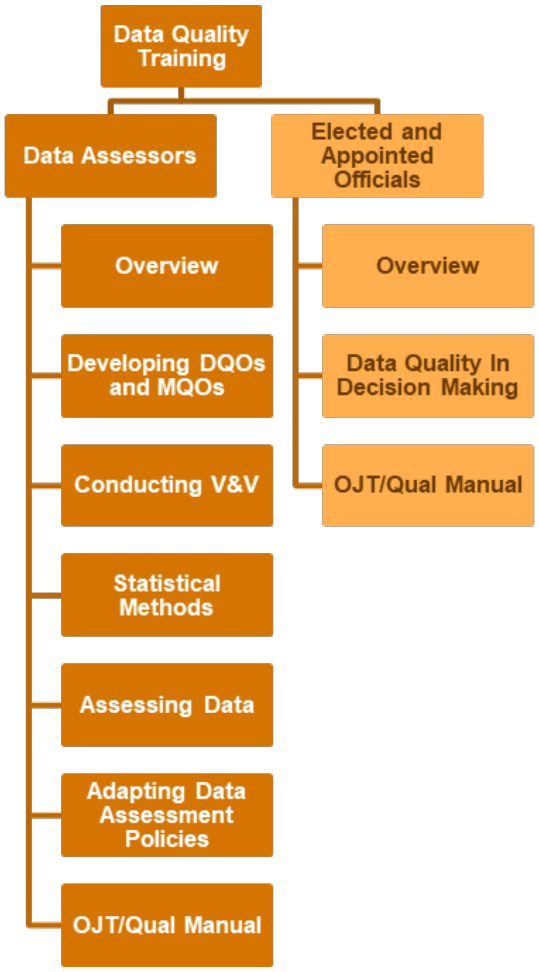


Figure 11. Example list of training and drill topics for responders and officials.

DQA Implementation.3	Guidance and regulations encyclopedia Develop a database that can be used to look up specific regulations and guidance pertinent to a specific response scenario that will inform decisions to be made and action levels that measurement data will need to be compared to.

A common resource for data assessors and response leadership would help establish a common understanding of the regulations that apply to a radiological incident, the reasoning behind them, and the specific action levels that need to be tested against. This understanding informs both immediate goals for release characterization and long-term goals for establishing the end of emergency operations and the transition to recovery, long-term monitoring, and remediation.

The guidance and regulations encyclopedia would be a database containing information on radiological protective action guidance and regulations in FSLTT jurisdictions. The database could be developed to allow users to search for protective action guidance and regulatory descriptions in several different ways, including text-based keyword searches, entering the

name of a location or U.S. state, or selecting a location on an interactive map. The descriptions would be presented in full, text-based form, or as simpler tables relating to specific action levels and a brief context statement. Action levels described by the regulations and guidance could then be selected to provide input to the assessment process. An example of a similar tool for emergency management guidance and regulations is the Interactive Requirements Document tool hosted on the NNSA Enterprise Data Management System.

DQA Implementation.4	Equipment data connections to CBRNResponder Develop incentives for radiation detection instruments to be manufactured with default built-in connection options to CBRNResponder.

Recording data manually can lead to high transcription error rates that can result in high rates of rejected data. Reducing the number of selections a user needs to enter to record data would decrease this error rate. Additionally, having mandatory input fields for a measurement would reduce the number of “Incomplete” and “Suspicious” flags on collected data. CBRNResponder is already popular with SLTT responders and is an ideal location to accumulate large amounts of data for a specific incident. Therefore, instruments should be able to upload data to CBRNResponder in an automated fashion. This would ideally be done by streaming the data as it is collected but including the ability to produce a CBRNResponder-compatible data file would also decrease error rates. Responders should be able to activate a feature on their instruments, and possibly other equipment such as mobile devices, to be able to upload measurements in an automated fashion such that data populates in the CBRNResponder event under the user’s organization. An example of this capability can be found in the Berkeley Nucleonics SAM 950 Radioisotope Identifier (BNC 2022).

For custom-developed instruments, modifications to the instrument software or firmware can be made by the agency responsible for developing the instrument. Commercial vendors cannot, of course, be required to implement such integration. However, requirements can be established by SLTT organizations that do not allow the purchase of equipment that is not able to upload data to CBRNResponder. Additional support could also be provided by federal agencies through programs like Securing the Cities to cover the costs of additional connections equipment and software required for some systems to connect. Such programs could also require that equipment purchased by funding from the program be required to have a CBRNResponder upload capability.

DQA Implementation.5	Integrated measurement procedure guide Develop a step-by-step procedure on instrument interfaces that guide users through measurement procedures according to an established QA/QC process.

In the fast-paced response environment, responders may sometimes focus too much on simply collecting the requested measurements and pay less attention to capturing the context for how the measurement was performed, which is important for data assessment. Manuals describing measurement data that need to be captured and QA/QC best practices are available but are seldom carried and referred to when performing measurements. If instead, the measurement procedure was built into the instrument user interface such that the user was prompted to follow

the QA/QC steps and record contextual information, all the information necessary for data assessment would be more likely to be captured and result in less rejected data.

For example, a step-by-step procedure could be implemented on the instrument interface that guides the user through measurement procedures found in the Radiological Dispersal Device (RDD) Response Guidance – Planning for the First 100 Minutes (DHS 2017). A more sophisticated functionality might also include dispatching measurement orders to personnel in the field through mobile or instrument applications to coordinate survey and sampling locations. Then, when the desired location was reached, the measurement procedure would be displayed. If a location could not be reached, the user would have the option to indicate the condition, and still collect a measurement.

An integrated measurement procedure guide would provide a stronger process-based approach for measurement collectors, as well as provide better tracking of initial data and contextual information for evaluation by data assessors. The more sophisticated option would provide a more strongly integrated approach to survey and sample planning to support DQOs. Implementation of this capability may be best supported through a partnership between Department of Energy (DOE) laboratories, instrumentation manufacturers, and responders who perform measurement collection. This partnership would see the DOE laboratories develop the integrated procedure guides and sample planning tasking. Industrial equipment manufacturers would perform the implementation on device interfaces. One or more pilot teams composed of first responders and other data collectors would then evaluate the features and provide feedback to laboratory and industry partners for additional development cycles.

DQA Implementation.6	Develop a background radiation level database Develop a database specifically to store background data collected by SLTT jurisdictions during normal operations as a reference for emergency operations.

Knowing the amount of regular background radiation at a location is important for determining whether, during a response to a radiological incident, that location has been contaminated. Oftentimes, background information has not been collected before an incident, and collecting background during the incident may not be appropriate if the area is now under suspicion of being contaminated.

For such circumstances where responders are unable to collect background measurements, having a database with applicable spectra or count rate values for the location from a variety of different instruments would help improve the ability of data assessors to evaluate the area for evidence of contamination. For example, historical background measurements available during the Portsmouth and Woolsey fire responses were valuable for comparisons to measurements taken during the response to establish that no additional contamination was present.

The authors recommend developing a national background database for gathering and making available background measurements from studies conducted throughout the United States. This concept has been attempted in various forms, including the DHS Background Catalog[1] and current DOE effort to gather information on current and former DOE facilities, but there is currently no commonly employed database used for radiological response that is available to all

[1] Department of Homeland Security, Countering Weapons of Mass Destruction. 2021. *Data Mining Analysis and Modeling Cell. Fiscal Year 2020 Annual Report.* Available from authors upon request.

responders. Possible sources of background measurements would include remediation documentation for former nuclear facility sites and published studies. Measurement data could also be collected from National Background Week events in CBRNResponder previously conducted in 2019 and 2021, as well as future National Background Week events, which could be held on a regular schedule to encourage measurements with new instruments and personnel. Conducting such events during normal operations would also provide more time to collect very accurate and precise background data to mitigate issues using such data during a radiological incident.

DQA Implementation.7	Develop guidance for decision makers Develop guidance explaining data quality and statistical concepts and how improving the quality of that data would translate into defensible, traceable, and robust decisions.

While regulations and guidance are available to assist officials making critical decisions that affect the lives, property, and livelihoods of the public, these decisions must be made quickly and defensibly. When a response agency prepares a data product for officials, those data products do not come with recommendations about what decision should be made based on the available data. It is left up to the official to synthesize multiple data products and other information inputs into decisions. For example, a data assessor might communicate that the data indicate “95% confidence that 95% of an area is below background.” However, it can be confusing for those who do not work daily with statistics to interpret the meaning of such a statement. Response leadership may ignore those values in favor of the purely visual aspects of data products, or else question those values and ask responders for 100% confidence without awareness of the implications of that request.

Additionally, it would be useful for federal agencies to evaluate data products currently being generated for their efficacy in communicating the pertinent information in an understandable way. A joint effort between technical subject matter experts and science communications specialists would be especially helpful in this endeavor to ensure that the data products currently being generated have maximum efficacy and impact to enhance response leadership’ understanding and help inform critical decisions.

7.0 Conclusions

V&V and DQA are critical to generating defensible conclusions based on information collected during a radiological incident. The project team gathered a summary of data assessment processes and methods through a literature review and additional research on V&V and DQA best practices to identify processes that may be useful during the early and intermediate phases of a radiation emergency incident. Additionally, this report was supplemented by outreach to partners at the local, state, and federal level while also incorporating operational lessons learned from other current and past efforts. Numerous standards and documents published by the U.S. Environmental Protection Agency, Federal Radiological Monitoring and Assessment Center, and as collaborative efforts between these groups and other organizations, were surveyed and investigated to collect these V&V and DQA requirements and procedures.

Current data assessment guidance is tailored specifically to a technical audience having knowledge of radiation health physics and statistics, and experience with the practical applications of radiation protection and in a non-emergency context. The methods described in the guidance often require relatively large datasets to form distributions such that assumptions hold for statistical testing. This requires ample time for data assessors to plot, analyze, and perform statistical testing, and the required documentation to provide defensible conclusions from these assessments. Therefore, as found, they are impractical for a majority of responders during the emergency response phase.

Recommendations are provided in this report to adapt current data assessment guidance for application to various response objectives in the development of the data assessment toolkit. Ultimately, the findings presented through this technical research aim to produce actionable, operational tools and guidance for responders. The authors have also presented additional recommendations regarding implementation of data assessment in responder tools and procedures for consideration beyond the scope of the DQA toolkit project.

This page intentionally left blank.

8.0 References

ANSI – American National Standard Institute. 2012. *Verification and Validation of Radiological Data for Use in Waste Management and Environmental Remediation*. ANSI/ANS-41.5-2012. https://webstore.ansi.org/standards/ansi/ansians412012

BNC – Berkeley Nuclionics Corp. 2022. "SAM 950 (RIID) Ruggedized Radiation Isotope Identifier." Accessed February 4, 2022 at https://www.berkeleynucleonics.com/sam-950-riid

Bowen, W.M. and C.A. Bennett (eds.). 1988. *Statistical Methods for Nuclear Material Management*. NUREG/CR-4604, PNL-5849. U.S. Nuclear Regulatory Commission, Washington, D.C. https://www.nrc.gov/docs/ML1116/ML11161A056.pdf

Buddemeier, B., D. Haynes, A. Wood-Zone, and G. Klemic. 2017. *Job Aids for Using Preventive Radiological/Nuclear Detection Equipment for Consequence Management*. LLNL-TR-739804. Lawrence Livermore National Laboratory, Livermore, California. https://www.osti.gov/servlets/purl/1400097

DHS – U.S. Department of Homeland Security. 2016. *Nuclear/Radiological Incident Annex to the Response and Recovery Federal Interagency Operational Plans.* U.S. Department of Homeland Security, Washington, D.C. https://remm.hhs.gov/NRIA_FINAL_110216.pdf

DHS – U.S. Department of Homeland Security. 2017. *Radiological Dispersal Device (RDD) Response Guidance, Planning for the First 100 Minutes*. Published by Department of Homeland Security: Science and Technology, National Urban Security Technology Laboratory, and Federal Emergency Management Agency; and the U.S. Department of Energy, National Nuclear Security Administration. https://www.dhs.gov/sites/default/files/publications/NUSTL_RDD-ResponsePlanningGuidance-Public_171215-508.pdf

EPA – U.S. Environmental Protection Agency. 1996. *The Volunteer Monitor's Guide To Quality Assurance Project Plans*. EPA 841 B-96-003. U.S. Environmental Protection Agency, Washington, D.C. https://www.epa.gov/sites/default/files/2015-06/documents/vol_qapp.pdf

EPA – U.S. Environmental Protection Agency. 2000. *Guidance for Data Quality Assessment: Practical Methods for Data Analysis. EPA QA/G-9, QA00 Update.* EPA/600/R-96/084. U.S. Environmental Protection Agency, Washington, D.C. https://www.epa.gov/sites/default/files/2015-06/documents/g9-final.pdf

EPA – U.S. Environmental Protection Agency. 2002. *Guidance on Environmental Data Verification and Data Validation*, EPA QA/G-8. EPA/240/R-02/004. U.S. Environmental Protection Agency, Washington, D.C. https://www.epa.gov/sites/default/files/2015-06/documents/g8-final.pdf

EPA – U.S. Environmental Protection Agency. 2006a. *Data Quality Assessment: Statistical Methods for Practitioners*, EPA QA/G-9S. EPA/240/B-06/003. U.S. Environmental Protection Agency, Washington, D.C. https://webcms-prod-en-uploads-20210415024638913800000003.s3.amazonaws.com/files/2015-08/documents/g9s-final.pdf

EPA – U.S. Environmental Protection Agency. 2006b. *Guidance on Systematic Planning Using the Data Quality Objectives Process*, EPA QA/G-4. EPA/240/B-06/001. Environmental Protection Agency, Washington, D.C. https://www.epa.gov/sites/default/files/2015-06/documents/g4-final.pdf

FEMA – Federal Emergency Management Agency. 2017. *National Incident Management System*. Federal Emergency Management Agency, Washington, D.C. https://www.fema.gov/sites/default/files/2020-07/fema_nims_doctrine-2017.pdf

FEMA - Federal Emergency Management Agency. 2022. "Radiological Operations Support Specialist." Department of Homeland Security. https://www.fema.gov/sites/default/files/2020-07/fema_cbrn-ross.pdf

Fournier, S.D., G. Cerefice, T. Laiche, and A. Moore. 2020. *RadResponder Nationwide Drill 2020: FRMAC Participation After Action Report*. SAND2020-2546R, Sandia National Laboratories. https://www.osti.gov/servlets/purl/1602952

FRMAC – Federal Radiological Monitoring and Assessment Center. 2010. *FRMAC Operations Manual*. DOE/NV/25946–980. U.S. Department of Energy, National Nuclear Security Administration, Nevada Site Office. https://www.nnss.gov/docs/docs_FRMAC/FRMAC%20Operations%20Manual%202010.pdf

Matzke, B.D., L.L. Newburn, J.E. Hathaway, L.M. Bramer, J.E. Wilson, S.T. Dowson, L.H. Sego, and B.A. Pulsipher. 2014. *Visual Sample Plan Version 7.0 User's Guide*. PNNL-23211. Pacific Northwest National Laboratory, Richland, Washington. https://vsp.pnnl.gov/docs/PNNL-23211.pdf

Millard, S.P. and N.K. Neerchal. 2001. *Environmental Statistics with S-Plus*. CRC Press, Boca Raton, Florida.

Multi-Agency. 2000. *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)*. NUREG-1575, Rev. 1; EPA 402-R-97-016, Rev. 1; DOE/EH-0624, Rev. 1. U.S. Nuclear Regulatory Commission, U.S. Environmental Protection Agency, and U.S. Department of Defense. https://www.epa.gov/sites/default/files/2017-09/documents/marssim_manual_rev1.pdf

Multi-Agency. 2004. *Multi-Agency Radiological Laboratory Analytical Protocols Manual (MARLAP): Assessment, Planning, Implementation*. NUREG-1576, EPA 402-B-04-001A, NTIS PB2004-105421. U.S. Environmental Protection Agency, U.S. Department of Defense, U.S. Department of Energy, U.S. Department of Homeland Security, U.S. Nuclear Regulatory Commission, Food and Drug Administration, U.S. Geologic Survey, National Institute of Standards and Technology. https://www.epa.gov/radiation/marlap-manual-and-supporting-documents

OriginLab. 2022. "Origin and OriginPro." Accessed February 4, 2022 at https://www.originlab.com/index.aspx?go=Products/Origin.

PNNL – Pacific Northwest National Laboratory. 2022a. "Nonparametric (Distribution-Free) Tolerance Limits." Accessed February 4, 2022. https://vsp.pnnl.gov/help/vsample/Design_UTL_Non_Parametric.htm

PNNL. 2022b. "Presence/Absence Compliance Sampling." Accessed February 4, 2022 at https://vsp.pnnl.gov/help/Vsample/Design_Building_Compliance_Sampling.htm

Schilling, E.G. and D.V. Neubauer. 2009. *Acceptance Sampling in Quality Control*. 2nd ed. CRC Press, Taylor & Francis Group, Boca Raton, Florida.

Squeglia, N.L. 2008. *Zero Acceptance Number Sampling Plans*. 5th ed. ASQ Quality Press, Milwaukee, Wisconsin.

This page intentionally left blank.

Appendix A – Glossary

Accuracy – "How close is each measurement to the true value?" This data parameter describes the comparison of a result to a consensus value, generally expressed in terms of an error, either as an absolute value or percentage, where the measurement is compared to a mean or known true value.

Action Level – A quantitative value that is compared to measured values to determine whether a specific set of actions is warranted. For example, an action level of 1 rem for evacuating a local population may be established based on the 2017 EPA Protective Action Guidance manual[1].

Assessment Phase – The third phase of the data life cycle, which occurs after data and information have been collected. The assessment phase includes the verification, validation, and data quality assessment processes. The output of the assessment phase is a set of conclusions about the data.

Comparability – "Can two measurements or sets of measurements be compared?" This data parameter describes whether measurements can be compared based on the instrument used, the type of measurement, the units used, and the question being answered. This is generally a qualitative parameter.

Completeness – "Did we collect all the data we planned for?" This data parameter describes how many measurements were collected compared to the number of planned measurements, generally expressed as a percentage of the actual to the planned numbers of measurements. It can also describe the completeness of the information collected for each individual measurement.

Data – The term for collections of both quantitative and qualitative facts and observations. For example, "9 mR/hr" and "Jane Smith" and "the corner of Cherry and 9th Ave." are data. Data requires context to be actionable, at which point it is transformed into information.

Data Life Cycle – The term for the data quality process composed of the planning, implementation, and assessment phases. The goal of the data life cycle is to support traceable and defensible conclusions based on data and information collected to address a question or meet an objective. The data life cycle addresses decision and estimation objectives, where the output is a conclusion about a specific parameter of the data, such whether the mean of the data is above or below a certain threshold, or an estimation of a statistical parameter with known accuracy and precision, such as 1.9 ± 0.2 mR/hr.

Data Quality – A collection of qualitative and quantitative information about a measurement or set of measurements that indicates whether or not that data can be used to address a specific question or objective or meet specified requirements for estimating a parameter of the data, such as the mean.

[1] U.S. Environmental Protection Agency. 2017. *PAG Manual: Protective Action Guides and Planning Guidance for Radiological Incidents*. EPA-400/R-17/001 https://www.epa.gov/sites/default/files/2017-01/documents/epa_pag_manual_final_revisions_01-11-2017_cover_disclaimer_8.pdf

Data Quality Assessment – The process of evaluating verified and validated data against the established data quality objectives for its suitability in addressing a decision or estimation objective. This process may involve performing statistical tests on groups of measurements, and results in a data assessor drawing conclusions about the data relative to the specific decision or estimation objective.

Data Quality Objectives (DQOs) – Qualitative and quantitative requirements on data and information necessary to support a specific decision or estimation objective. DQOs are established in the planning phase.

Implementation Phase – The second phase of the data life cycle, which occurs as data and information are being collected. The implementation phase includes quality assurance and quality control practices and processes. The output of the implementation phase is data and information that meet the data quality objectives established in the planning phase of the data life cycle.

Information – The term for collections of data that have been organized into a meaningful and useful context. For example, the data, “9 mR/hr” and “Jane Smith” and “the corner of Cherry and 9th Ave.” can be transformed into information by tying them together: “Jane Smith measured 9 mR/hr at the corner of Cherry and 9th Ave”.

Measurement – The term for the act of quantifying physical phenomena. In this report, the term measurement refers to the information generated by the act of quantifying physical phenomena, specifically radiological phenomena.

Measurement Quality Objectives (MQOs) - Qualitative and quantitative performance requirements for measurement data and information characteristics, such as precision and accuracy, based on established data quality objectives.

Planning Phase – The first phase of the data life cycle, which occurs before data and information are collected. The planning phase includes the data quality objectives process and generation of measurement quality objectives. The output of the planning phase is a set of data quality objectives and measurement quality objectives.

Precision – “How much variation is there between measurements?” This data parameter describes the reliability by which an investigator can reproduce the sample results. It measures the amount of dispersion among series of measurements and is often provided as a standard deviation.

Quality Assurance – Encompasses all of the actions necessary to provide confidence that the data and information collected during a measurement are of sufficient quality to be used to support a specific decision or estimation objective. This includes recording information about the circumstances of a measurement, such as the instrument calibration. Quality assurance is mainly conducted in the implementation phase.

Quality Control – Encompasses all of the actions that control and measure the circumstances of a measurement. For example, recording the height of a measurement above the ground is a quality control action. Quality control is mainly conducted in the implementation phase.

Representativeness – “Do the samples come from the same area?” This data parameter describes how well a sample characterizes or describes a specific population and is generally

reported as a qualitative basis. This can be evaluated, for example, by checking whether measurements were taken on the same kind of terrain, or within a certain distance from each other.

Sensitivity – “Are the instruments used sensitive enough to detect what we are looking for?” This data parameter describes at what threshold value an instrument can detect radiation. Instruments may have several different sensitivity thresholds for different radiation types and for different specific radioisotopes.

Validation – The act of comparing measurement information collected against a list of measurement quality objectives and noting whether the measurement quality objectives have been met. Data validation asks the question, “Were the measurement quality objectives achieved for each measurement and as a whole?” The purpose of data validation is to determine if the measurement parameter requirements for the parameters have been achieved by the measurements being considered

Verification – The act of comparing measurement information collected against a list of measurement information required and noting whether the information collection requirements have been fulfilled. Verification asks the question, “Did all of the requested information get collected?” The purpose of data verification is to ensure that the records associated with a specific set of radiological measurements reflect all of the processes and procedures used to generate that dataset.

Appendix B – Data Assessment Process Example

Objective: Assess Reports of Radiological Contamination

Background: There has been an explosion in downtown Anytown, USA, on January 1, 2022, at 10:00 am, and it is suspected that the device may have been a radiological dispersion device (RDD). Reports and radiation measurements have been transmitted to the emergency operations center (EOC). A data assessor has been assigned to determine whether any elevated radiation levels are present (greater than three times background).

Collect Data: The data assessor collects or gains access to all the reports they can obtain about the incident so far. The data are collected as shown in Table B.1.

Table B.1. Data collected from all incident reports

Data	Measurement 1	Measurement 2
Name of the individual	Jane Smith	John Smith
Organization of the individual (if applicable)	Anytown Fire Department	Anytown Police Department
Location	Corner of Cherry and 9th Ave.	47.606, -122.326
Orientation of instrument to nearby surfaces		
Distance from instrument to nearby surfaces		
Type of material on nearby surfaces relative to the detector		
Time of day and date the measurement was taken	10:05 am, January 1, 2022	January 1, 2022
Instrument make and model, and of associated probes		BNC NucALERT
Instrument serial number		
Instrument calibration or certification records		
Conversion factor from counts per minute (cpm) to disintegrations per minute (dpm) for the instrument (if applicable)		
Instrument settings		
Survey method – stationary or scanning (an instrument in a moving vehicle is considered a scan)	Walking	Walking
If the survey meter was in motion, the speed of the meter (speed of the vehicle or approximate walking speed)		
Background measurement		
Measurement value	9	2
Measurement unit	mR/hr	
Distance between each successive measurement		
Alarm type (if applicable)	Gamma	Gamma
Isotopes identified (if applicable)	n/a	n/a
Isotope identification confidence (if applicable)	n/a	n/a
Shipping manifest (if applicable)	n/a	n/a

Verification: The data assessor applies flags to the data based on the verification table (Table 1, Section 4.1) to each measurement individually. The flags are highlighted in red to show they have been applied (Table B.2).

Table B.2. Flags applied to data based on verification

Data	Measurement 1	Measurement 2
Name of the individual	Jane Smith	John Smith
Organization of the individual (if applicable)	Anytown Fire Department	Anytown Police Department
Location	Corner of Cherry and 9th Ave	47.606, -122.326
Orientation of instrument to nearby surfaces	**Suspicious**	**Suspicious**
Distance from instrument to nearby surfaces	**Suspicious**	**Suspicious**
Type of material on nearby surfaces relative to the detector		
Time of day and date the measurement was taken	10:05 am, January 1, 2022	January 1, 2022 **Suspicious – no time**
Instrument make and model, and of associated probes	**Incomplete**	BNC NucALERT
Instrument serial number		
Instrument calibration or certification records		
Conversion factor from counts per minute (cpm) to disintegrations per minute (dpm) for the instrument (if applicable)		
Instrument settings		
Survey method – stationary or scanning (an instrument in a moving vehicle is considered a scan)	Walking	Walking
If a moving survey method was used, the approximate speed of the meter or vehicle the meter was in	**Suspicious – no walking speed**	**Suspicious – no walking speed**
Background measurement	**Suspicious**	**Suspicious**
Measurement value	9	2
Measurement unit	mR/hr	
Alarm type (if applicable)		
Isotopes identified (if applicable)		
Isotope identification confidence (if applicable)		
Shipping manifest (if applicable)		

Validation: The data assessor applies flags to the measurements based on the on the validation table (Table 2, Section 4.2). The flags are highlighted in red to show they have been applied (Table B.3).

Table B.3. Flags applied to data based on validation

Data	Measurement 1	Measurement 2
The instrument does not provide a dose readout, but value representative of a dose range instead.	**none**	**Imprecise**
1. The calibration expiration date is earlier than the date the measurement was taken. 2. No calibration information is available.	**none** – no calibration information provided	**none** – no calibration information provided
The measured value is in the lower 20% of the full-scale range.	**none**	**none**
A given measurement is less than three times the typical background for the area with the type of instrument used.	**none** – no background measurement	**none** – no background measurement
The measurement is consistent with an elevated background, and any identified isotopes are consistent with NORM.	**none** – no isotopes	**none** – no isotopes
Five or more ”Suspicious” flags have been applied.	**none**	**Suspicious Context**
All measurements are within 15 meters (50 feet) of one another.	**none** – using an online map, the data assessor determines that the distance between the measurements is approximately 303 ft.	

Data Quality Assessment: The data assessor performs a quality assessment and draws conclusions based on the quality assessment guidance (Table 3, Section 5.1). Based on this guidance, the data assessor makes notes about the measurement set as a whole. The red highlights show where the condition on the left has been met, and therefore indicates an issue with the measurement set (Table B.4).

Table B.4. Measurement set notes after data quality assessment

Condition	Measurement Set
The “Imprecise” flag has been applied to an individual measurement.	**Measurement 2 has the Imprecise flag – the NucALERT displays a number representing a dose range.**
The “Out of Calibration” flag has been applied.	No measurements with the Out of Calibration flag.
The “Invalid” flag has been applied to an individual measurement.	No measurements with the Invalid flag.
Every measurement in the dataset has had the “Background” or “NORM” flag applied to it.	No measurements with the Background flag.
The “Incomplete” flag has been applied to an individual measurement.	**Measurement 1 has the Incomplete flag – Jane Smith did not include the instrument make and model used to make the measurement.**
The “Suspicious Context” flag has been applied to a measurement.	**Measurement 2 has the “Suspicious Context” flag.**

Condition	Measurement Set
The number of ”Suspicious” flags for a set of measurements is equal to or greater than three times the number of measurements.	**There are 9 total “Suspicious” flags, which is greater than 6 (3 * 2 measurements).**
The number of measurements that do not have “Incomplete” or “Invalid” flags is less than two.	**The number of measurements that do not have “Incomplete” or “Invalid” flags is one.**
Only one instrument was used to make all valid measurements.	Two different instruments were used to perform the measurements.
None of the above conditions is true.	Some of the above conditions are true.

The data assessor returns to the validation step and takes the following actions:

1. The measurement from Jane Smith is not usable as is since the instrument make and model are missing. The data assessor contacts Jane to get the instrument information and is successful – Jane was using a Polimaster PM1703GN, which is appropriate for radiation surveys at less than 10 mR/hr, and the data assessor clears the “Incomplete” flag. The data assessor also asks about the distance from the instrument to the ground and its orientation to other surfaces, but Jane does not recall that information, so the data assessor leaves the “Suspicious” flags in place. Finally, the data assessor assumes an average walking speed of 4 mph but does not clear the “Suspicious” flag. The data assessor determines that Jane’s measurement is now usable for this assessment.

2. The data assessor determines that the measurement from John Smith is usable for this response objective because only an indication of above-average radiation exposure levels is necessary. The data assessor uses the “Job Aids for Using Preventive Radiological/Nuclear Detection Equipment for Consequence Management” guidance to determine that John’s instrument was reading between 0.035 mrem/hr and 0.40 mrem/hr. There are no background measurements to compare to, but the data assessor knows that an average background reading for Anytown is approximately 0.02 mrem/hr at most, so John’s measurement does not exceed three times this number (0.06 mrem/hr). The data assessor determines that John’s reading does not indicate an increased radiation exposure level and applies a “Background” flag to the measurement. The data assessor also applies the “Suspicious” flag, since they are unable to contact John at this time, and John did not provide a background measurement with his instrument.

Table B.5 is the resulting verification table for the two measurements, Table B.6 is the resulting validation table, and the data assessor’s notes on the DQA table appear in Table B.7.

Table B.5. Verification table for two measurements

Data	Measurement 1	Measurement 2
Name of the individual	Jane Smith	John Smith
Organization of the individual (if applicable)	Anytown Fire Department	Anytown Police Department
Location	Corner of Cherry and 9th Ave	47.606, -122.326
Orientation of instrument to nearby surfaces	**Suspicious**	**Suspicious**
Distance from instrument to nearby surfaces	**Suspicious**	**Suspicious**
Type of material on nearby surfaces relative to the detector		
Time of day and date the measurement was taken	10:05 am, January 1, 2022	January 1, 2022 **Suspicious – no time**
Instrument make and model, and of associated probes	~~**Incomplete**~~ Polimaster PM1703GN	BNC NucALERT
Instrument serial number		
Instrument calibration or certification records		
Conversion factor from counts per minute (cpm) to disintegrations per minute (dpm) for the instrument (if applicable)		
Instrument settings		
Survey method – stationary or scanning (an instrument in a moving vehicle is considered a scan)	Walking	Walking
If a moving survey method was used, the approximate speed of the meter or vehicle the meter was in	**Suspicious – no walking speed** Assume 4 mph	**Suspicious – no walking speed** Assume 4 mph
Background measurement	**Suspicious**	**Suspicious**
Measurement value	9	2
Measurement unit	mR/hr	
Alarm type (if applicable)		
Isotopes identified (if applicable)		
Isotope identification confidence (if applicable)		
Shipping manifest (if applicable)		

Table B.6. Validation table for the two measurements.

Data	Measurement 1	Measurement 2
The instrument does not provide a dose readout, but value representative of a dose range instead.	**none**	~~**Imprecise**~~ **Suspicious** – Imprecise flag and no background measurement
1. The calibration expiration date is earlier than the date the measurement was taken. 2. No calibration information is available.	**none** – no calibration information provided	**none** – no calibration information provided
The measured value is in the lower 20% of the full-scale range.	**none**	**none**
A given measurement is less than three times the typical background for the area with the type of instrument used.	**none** – no background measurement. Assume background of 0.02 mrem/hr, threshold is 0.06 mrem/hr. This measurement is greater than three times the background.	**Background** ~~**none** – no background measurement.~~ Assume background of 0.02 mrem/hr, threshold is 0.06 mrem/hr. This measurement is less than three times the background.
The measurement is consistent with an elevated background, and any identified isotopes are consistent with NORM.	**none** – no isotopes	**none** – no isotopes
Five or more ”Suspicious” flags have been applied.	**none**	**Suspicious Context**
All measurements are within 15 meters (50 feet) of one another.	**none** – using an online map, the data assessor determines that the distance between the measurements is approximately 303 ft.	

Table B.7. Data assessor's notes on the DQA table

Condition	Measurement Set
The "Imprecise" flag has been applied to an individual measurement.	~~**Measurement 2 has the Imprecise flag – the NucALERT displays a number representing a dose range.**~~ Imprecise flag has been converted to a Suspicious flag for Measurement 2.
The "Out of Calibration" flag has been applied.	No measurements with the Out of Calibration flag.
The "Invalid" flag has been applied to an individual measurement.	No measurements with the Invalid flag.
Every measurement in the dataset has had the "Background" or "NORM" flag applied to it.	No measurements with the Background flag.
The "Incomplete" flag has been applied to an individual measurement.	~~**Measurement 1 has the Incomplete flag – Jane Smith did not include the instrument make and model used to make the measurement.**~~ No measurements with Incomplete flags.
The "Suspicious Context" flag has been applied to a measurement.	**Measurement 2 has the "Suspicious Context" flag.**
The number of "Suspicious" flags for a set of measurements is equal to or greater than three times the number of measurements.	**There are 9 total "Suspicious" flags, which is greater than 6 (3 * 2 measurements).**
The number of measurements that do not have "Incomplete" or "Invalid" flags is less than two.	~~**The number of measurements that do not have "Incomplete" or "Invalid" flags is one.**~~ Two measurements without Incomplete or Invalid flags.
Only one instrument was used to make all valid measurements.	Two different instruments were used to perform the measurements.
None of the above conditions is true.	Some of the above conditions are true.

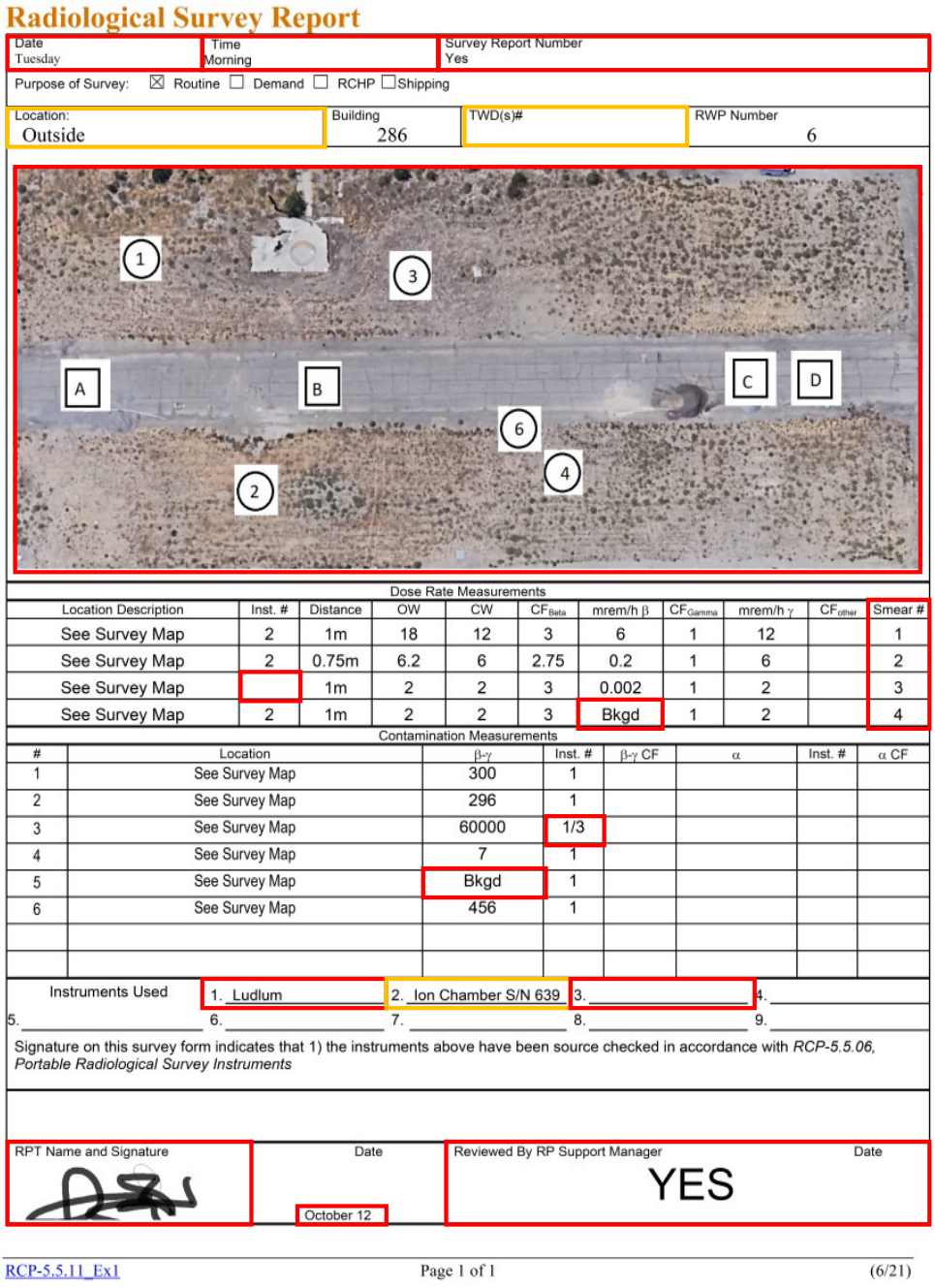
The data assessor determines that both measurements are now usable. The data assessor further notes that one measurement shows elevated radiation levels with certainty (Measurement 1), and that one that does not (Measurement 2, which has the "Background" flag applied). In addition, there are still a high number of "Suspicious" flags. The data assessor determines that additional measurements near Jane's position are needed to confirm the presence of radioactive contamination.

At this point, the data assessor might recommend that the Operations Section assign John to take a reading with his instrument near Jane's position (but still 50 ft away, toward the explosion, if possible). Alternatively, other measurements may have been reported during the time the data were being assessed, and the data assessor might first assess those to see if they are suitable for this response objective. In any case, the data assessor should apply the verification and validation steps to any new measurements that are collected, and then consider the new measurements and the original measurements together again for the quality assessment step.

The data assessor may find, upon reviewing other measurements, that there does appear to be radioactive contamination at the scene. Now the data assessor has been assigned to determine the radiation exposure level at the scene and start to assign zones per the "Radiological Dispersal Device (RDD) Response Guidance: First 100 Minutes Guidance". This is a different objective than the one just completed and has different data quality requirements. The data assessor's new objective is to determine if there are dose rates above a certain action level. Now, the data assessor must collect any new data available and begin reviewing the data again under the requirements for the new objective. In this case, John's measurements with his BNC NucALERT are not appropriate since his instrument does not give a dose reading directly, and so new measurements may need to be collected.

Appendix C – Example Incorrectly Completed Sample Form

Figure C.1 shows an example survey and sample form that has been incorrectly filled out in several places. This is an example based on a PNNL internal survey report form. No units are listed here because a standard for recording units is in place internally. This is an example form only and not based on a real survey.



Radiological Survey Report

Date	Time	Survey Report Number
Tuesday	Morning	Yes

Purpose of Survey: ☒ Routine ☐ Demand ☐ RCHP ☐ Shipping

Location:	Building	TWD(s)#	RWP Number
Outside	286		6

Dose Rate Measurements

Location Description	Inst. #	Distance	OW	CW	CF_{Beta}	mrem/h β	CF_{Gamma}	mrem/h γ	CF_{other}	Smear #
See Survey Map	2	1m	18	12	3	6	1	12		1
See Survey Map	2	0.75m	6.2	6	2.75	0.2	1	6		2
See Survey Map		1m	2	2	3	0.002	1	2		3
See Survey Map	2	1m	2	2	3	Bkgd	1	2		4

Contamination Measurements

#	Location	β-γ	Inst. #	β-γ CF	α	Inst. #	α CF
1	See Survey Map	300	1				
2	See Survey Map	296	1				
3	See Survey Map	60000	1/3				
4	See Survey Map	7	1				
5	See Survey Map	Bkgd	1				
6	See Survey Map	456	1				

Instruments Used: 1. Ludlum 2. Ion Chamber S/N 639 3. _____ 4. _____ 5. _____ 6. _____ 7. _____ 8. _____ 9. _____

Signature on this survey form indicates that 1) the instruments above have been source checked in accordance with *RCP-5.5.06, Portable Radiological Survey Instruments*

RPT Name and Signature	Date	Reviewed By RP Support Manager	Date
	October 12	YES	

RCP-5.5.11_Ex1 Page 1 of 1 (6/21)

Figure C.1. Example of a survey and sample form that has been incorrectly completed in several places (red boxes), with suspicious information in some places (yellow boxes).

The parts if the form in Figure C.1 that are incorrect or incomplete are described as follows. The issues correspond to the red and yellow boxes in the figure.

- The Date reads “Tuesday.” It should indicate the day of the month, the month, and the year the surveys and samples were conducted (e.g., October 12, 2021).
- The Time reads “Morning.” It should list the time of day, and ideally also the time zone (e.g., 10:45 a.m. Pacific).
- The Survey Report Number reads “Yes.” This should be an official document number for tracking.
- The location reads “Outside.” While the building number given is helpful, a more detailed description would be better (e.g., “Outside, west of the building”).
- The technical work document (TWD) number is blank. Although this is an internal control, it would be better if this were filled in because it would describe how the surveys were conducted and give a data assessor a better idea of what methods were used to derive the values in the form.
- The instrument number (Inst. #) in line 3 of the Dose Rate Measurements section is blank. This should be filled in so that a data assessor knows which survey it relates to.
- The mrem/h β in line 4 of the Dose Rate Measurements section reads “Bkgnd.” This is not helpful since it does not list a specific value. A quantitative value should be reported (e.g., 2).
- The Smear Number (Smear #) in the Dose Rate Measurements section lists 1, 2, 3, and 4. These should be dose rate measurements A, B, C, and D instead.
- The instrument number (Inst. #) in line 3 of the Contamination Measurements section reads “1/3.” This seems to indicate that 60,000 cpm was measured on two different instruments, which is unlikely, given the random nature of radioactive decay. If the spot was measured by both instruments 1 and 3, then two values should be listed, or a note inserted indicating the value is an average. If instrument 3 measures a different quantity than instrument 2, then other information in this line needs to be filled out.
- The β-γ in line 5 of the Contamination Measurements section reads “Bkgnd.” This is not helpful since it does not list a specific value. A quantitative value should be reported, (e.g., 10).
- Instrument 1 in the Instruments Used section reads “Ludlum.” This should also list a model number and serial number to identify the specific instrument used (e.g., Ludlum Model 6 Geiger Counter S/N 123 with Model 44-9 Probe S/N 456).
- Instrument 2 in the Instruments Used section reads “Ion Chamber S/N 639.” Ideally the make and model would also be listed here. However, the make and model can be referenced using the serial number (S/N), so this entry is acceptable.
- Instrument 3 is blank in the Instruments Used section. Because the Contamination Measurements section indicates that instrument 3 was used, this line should be filled in (e.g., Ludlum Model 6 Geiger Counter S/N 789 with Model 44-9 Probe S/N 101).
- The RPT Name and Signature section includes only a signature. This section should also include the radiation protection technologist’s printed name (e.g., John Smith).
- The Date in the RPT Name and Signature section reads “October 12.” This should also include the year (e.g., October 12, 2021).

- The Reviewed by RP Support Manager section reads “Yes.” This section needs to include the name of the RP Support Manager, their signature, and the date.

Pacific Northwest National Laboratory

902 Battelle Boulevard
P.O. Box 999
Richland, WA 99354
1-888-375-PNNL (7665)

www.pnnl.gov