Characterization Survey Plan
Piqua Nuclear Power Facility

Auxano

March 2020

Stephen L. Bump, CHP, CIH, PMP
Clark B. Barton, CHP

Prepared for:

Auxano
Lebanon, Ohio 45036

NV5

1835 Terminal Drive
Suite 200
Richland, WA 99354

DMA-TR-107

LIMITATIONS

NV5/Dade Moeller has prepared this survey plan based on historical documents and drawings supplied by the current Piqua project team, and information received from the current Piqua project team regarding historical work with radioactive material at the facility. Our work is based on information available at the time of publication. Information provided in this report by NV5/Dade Moeller is intended exclusively for the use of the Piqua project team. The characterization survey plan requirements discussed in this report are based on the information provided, and our current understanding and interpretation of regulatory agency regulations, guidance and policies. The professional services have been performed in accordance with practices generally accepted by other health physicists practicing in this field. No other warranty, either expressed or implied, is made.

EXECUTIVE SUMMARY

This survey plan is intended to provide the instructions to the surveyor(s) completing the characterization survey of the Piqua Nuclear Power Facility. A survey report will be developed, reviewed, and approved to document the results of the survey and any follow-on actions if required.

TABLE OF CONTENTS

RECORD OF REVISIONS

Revision Number	Description	Date
0	New Document issued by Clark Barton	03/14/2020

1.0 INTRODUCTION

This survey plan will be used to perform a characterization survey of the Piqua Nuclear Power Facility (PNPF). The PNPF consists of the Reactor Building, the Auxiliary Building, and the surrounding land of approximately ½ acre. The floor plans, initial classifications, and proposed survey locations for the buildings and land are shown in Attachments 1 through 5.

A Historical Site Assessment (HSA) (NV5/DM, 2020) for the PNPF shows that the entire Reactor Building encompassed work with radioactive material. Additionally, the Auxiliary Building basement, any piping that may have contained steam generated using coolant from the reactor, tanks that may have contained reactor related fluids, and the airlocks from the Reactor Building into the Auxiliary building, would be included in the Class 1 areas. Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)-like Class 1 surveys will be conducted for both alpha and beta-emitting materials in these areas. 100% of accessible surfaces will be scanned, the areas will be broken into nominally 100 square meter areas, and the required number of sample locations identified by the MARSSIM process will have direct static and removable contamination measurements performed.

The areas immediately adjacent to the airlock entries, the stairwells leading up from the auxiliary building basement, as well as any areas identified that contain steam pipes that once contained steam generated using reactor coolant, are included in the Class 2 areas. Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)-like Class 2 surveys will be conducted for both alpha and beta-emitting materials in these areas. A minimum of 75% of the accessible surfaces will be scanned, the area will be broken into nominally 500 square meter areas, and the required number of sample locations identified by the MARSSIM process will have direct static and removable contamination measurements performed.

The remainder of the Auxiliary Building was office and control operations and it is unlikely any work with radioactive material other than storing sealed check type source(s) was ever performed in the facility. There was no indication that contamination was ever found in the Auxiliary Building. In addition, the external surfaces of the buildings, as well as the surface of the ground have no historical evidence of contamination. Based on the lack of dispersible material, the Auxiliary Building, other than the basement, as well as the outside of the buildings and the surface of the ground, could be considered non-impacted. However, it was deemed prudent to survey portions of these areas. Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)-like Class 3 surveys will be conducted for both alpha and beta-emitting materials in these areas. In these areas, required number of sample locations identified by the MARSSIM process will have 1 square meter areas surveyed, with direct static and removable contamination measurements performed. The floors will be scan surveyed with the floor monitor on all accessible surfaces, and a gamma survey will be performed in all areas.

2.0 CHARACTERIZATION SURVEY REQUIREMENTS

For purposes of this survey, the appropriate release criteria will be removable contamination below 1000 dpm / 100 cm^2, average total contamination below 5000 dpm / 100 cm^2, and the maximum total activity on any 100 cm^2 area below 15,000 dpm / 100 cm^2. These criteria are generally acceptable for unrestricted release for DOE O 458.1, as found in 10CFR835 Appendix D and

MARSSIM (NRC, 2000) for the NRC/State of Ohio. Additionally, the data collected for this survey will be designed such that it could be used as all or part of a final status survey.

2.1 MINIMUM DETECTABLE ACTIVITY DETERMINATION

Minimum Detectable Activity (MDA) will be calculated to verify that the chosen survey instruments and methods are adequately sensitive to meet the survey criteria. A Decision Level (DL), also the DCGL, will be set to the release limits and used to determine whether a survey point meets the criteria. An Action Level (AL) will be calculated based on the background study. The AL will be set to 25% of the release limits.

In order to calculate the MDA, an acceptable false positive and false negative rate must be determined. For purposes of this survey, it is assumed that a 5% false positive (α) and a 5% false negative (β) rate will be acceptable. The number of standard deviations above zero on the standard normal distribution to have a probability of α or β being higher is the standard normal deviate, kα or kβ. From Table 5.2 of the MARSSIM (NRC 2000), values of kα and kβ are set at 1.645 when α and β are equal to 0.05 (5%).

Direct measurements are taken by placing a detector as near as possible to the surface measured. A discrete measurement is then performed for a pre-determined time interval and the reading recorded. Direct measurements are collected using data-logging instruments such as the Ludlum 2360 in combination with the Ludlum 43-93 detector. The instruments are calibrated with radionuclides of similar energies to the radionuclides of concern as discussed in Section 3.0.

Direct measurement MDAs for alpha and beta-gamma-emitting radionuclides are calculated using the following expanded equation based on MARSSIM Equation 6-7 (NRC, 2000):

$$MDA_{static} = \frac{3 + 2k\sqrt{R_b t_g (1 + {}^{t_g}/_{t_b})}}{t_g \cdot E_{total} \cdot \frac{A}{100}}$$

where:

MDA_{static}	= MDA level in dpm/100cm^2
R_b	= background count rate in cpm
t_b	= Background counting time
t_g	= sample counting time
E_{total}	= total detector efficiency for radionuclide emission of interest (includes combination of instrument efficiency and surface efficiency*)
A	= physical detector area in cm^2

* Surface efficiency applied is in accordance with the recommended values of 0.25 for alpha and beta emitters <400 keV max and 0.5 for all others as described in ISO-7503-1.

Wipe samples are performed to determine the presence of removable residual radioactivity at each location for which a direct measurement is made. A wipe is performed by wiping a standard wipe [47 mm-diameter, paper or absorbent cloth] over an area of 100 cm^2 using moderate pressure. Wipes are collected from each direct measurement location. Analysis for gross beta-gamma and alpha radioactivity are performed onsite by NV5/Dade Moeller.

The MDAs for wipe samples are determined using the same equation specified for direct measurements.

2.2 BUILDING AND MATERIAL SPECIFIC BACKGROUND

Building-specific background will be established using data analysis of the first surveys performed in the auxiliary building. Since the reactor and auxiliary buildings were built at the same time, there is no evidence of any contamination in the ground floor or above of the auxiliary building, and concrete floors and walls predominate in the reactor building, it is reasonable to use the auxiliary building above grade floors to establish background levels for the materials.

2.3 RADIONUCLIDES OF CONCERN

The HSA (NV5/DM 2020) and referenced documents list the radionuclides of concern for the PNPF, that still exceed 50 millicuries of activity (plus Ag-108m based on potential dose) in 2019 as follows: ^{60}Co, ^{3}H, ^{55}Fe, ^{63}Ni, ^{41}Ca and ^{108m}Ag. These are all beta-gamma emitters or electron capture decays. It was noted in the reports when the fuel was removed that examination of the fuel cladding showed no issues with cladding failures. There was some indication of potential minor uranium contamination in the cladding alloy, so alpha contamination is also a possibility.

Based on the available information, the screening for radionuclides of concern for PNPF will be limited to gross alpha and gross beta. This is appropriate given that none of the listed isotopes would be present in the background in building structural materials in levels detectable using methods proposed for this survey. Based on the ratios of the potential contaminates, with Co-60 being 27% of the activity, and the hard to detect (HTD) isotopes ^{3}H and ^{63}Ni being 69%, and the DCGL being set at 25% of the release limits, the survey will be sufficient to show the total activity is below the limits. Should a removable residual radioactivity sample exceed the DCGL, additional evaluation will be conducted including holding the sample for recount to determine if short lived radon daughter products are present and if a contaminated area is identified that exceeds the DCGL with long lived activity, an additional set of swipe samples will be collected for liquid scintillation counting to quantify the HTD isotope activity.

2.4 ACTION LEVELS

The action levels for PNPF will be established at a level that is 25% of the release criteria. If an action level is exceeded, then characterization activities will take place to determine the nature and extent of the area of elevated radioactivity and if it exceeds the release criteria.

2.5 DEVIATIONS FROM THE SAMPLE DESIGN

During execution of sample designs, obstructions sometimes prevent personnel from collecting a measurement or sample at the precise location described in the sample planning documents, resulting in a non-representative or unobtainable sample. In such cases, it may be appropriate to slightly alter the location of sample collection, rather than proceeding to the next sample location. During execution of the characterization survey, it is acceptable to move a sample location in order to obtain a representative sample, if the movement is recorded on sample documentation. If a representative sample is unobtainable within a reasonable limit (less than 2 meters), then the location should be marked as unobtainable and personnel should proceed to the next sample location. In addition, any identified areas of elevated activity will be added as bias sample points, as will all floor drains, unless the areas coincide with a preselect survey point.

2.6 SURVEY UNIT DEFINITION

The majority of the survey units are shown in Attachment 1. There are a total of 46 survey units for PNPF: 34 Class 1 units, 3 Class 2 Units, and 9 Class 3 units. Survey designs have been developed using Visual Sample Plan (PNNL, 2017) and the statistical tools described above for all of the surveys except for the 2 stairwells in the Reactor Building , 3 wall survey units in the Auxiliary Building, and any additional surveys that may be needed based on steam pipes, potentially contaminated vaults or tanks identified as project progress occurs.

For each survey unit a predetermined number of sample locations will be determined, based on the MARSSIM methodology. The calculated sample size is increased automatically by 20% by the VSP (PNNL 2017) program to account for the possibility of unusable sample locations, while maintaining the required statistical power of the data set. For example, if 12 areas are required, VSP will increase the number to 15. Due to geometry effects, the actual number of samples may be larger than 15 to ensure adequate area coverage.

Not every survey unit has a pre-printed map and survey point pre-selection. Class 3 wall surveys in the Auxiliary Building, Levels 100, 111, and 121, as well as Class 1 surveys of the Spiral and Normal Staircases in the Reactor Building, do not have maps and preselected survey points. All of these surveys will be documented on-site, with bias selected locations.

The Spiral Staircase cannot be used; therefore the surveyor will only be able to survey locations that can be reached from a stable, allowable location. 100% of all reachable areas will be scanned and at least 15 sample points will be selected for static and removable measurements.

The Normal Stairwell will be surveyed at the landings, floors, first two meters of the walls, and each stair tread. At least 15 sample points, up to 30, will be selected for static and removable measurements.

The first 2 meters of the walls in the Auxiliary Building Floors 100, 111, and 121 levels will be surveyed each as a survey unit. 30 sample points will be selected, with a scan of each full square meter surrounding the survey point. The sample points will be biased by selecting cracked, stained, or otherwise suspect areas when possible, with the points randomly scattered such that each room or hallway has at least 1 point.

Additional surveys for structures identified as the project progresses, such as potentially contaminated steam lines, vaults or tanks, will be added using the same parameters as the predesigned surveys discsussed in this document. The classification will be in accordance with the HSA and data collected as part of this survey.

2.7 SURVEY INSTRUMENTATION

The survey instrumentation for this survey plan includes:

- Ludlum Model 43-93, Dual Phosphor Scintillator, used for direct and scan measurements for alpha and beta residual radioactivity.
- Ludlum Model 239-1F, Gas proportional floor monitor, used for scanning survey of floor surfaces for alpha and beta residual radioactivity.
- Ludlum Model 43-10-1, Dual Phosphor Scintillator, used to analyze wipe samples for alpha and beta removable residual radioactivity.
- Ludlum Model 44-10, 2x2 NaI Detector, used to perform gamma scanning surveys.

Based on the potential radionuclides assumed to be present and their associated radiations, the detection sensitivities of various instruments and techniques were evaluated for use at the PNPF. Instruments are selected for use during direct measurements and analysis of removable residual radioactivity wipes. The instruments were selected to meet the Minimum Detectable Activity (MDA) goals. Instrument selection criteria include the capability of the instrument to detect the radiation types emitted by the radionuclides of concern, sufficient efficiency factors, and estimates of background radiation values.

Instruments are calibrated with radionuclides of similar energies to the radionuclides of concern and are traceable to the National Institute of Standards and Technology (NIST). Since the release limits are lower for removable residual radioactivity, the wipes will be counted on equipment that is more sensitive than the field instruments used for the direct measurements. Table 2-1 shows proposed instrumentation. It is noted that equivalent instrumentation, approved in advance, is also acceptable.

Table 2-1: Proposed Instrumentation

Detector Model	Detector Type	Detector Area	Instrument Model	Window Thickness	Total Efficiency
Ludlum 43-93	Dual Phosphor Scintillation	100 cm^2	Ludlum 2360	1.2 mg/cm^2	7.5% - Beta[1] 10% - Alpha[1]
Ludlum 43-10-1	Dual Phosphor Scintillation	N/A	Ludlum 3030E or 2929	0.8 mg/cm^2	27% – Beta 4π 39% – Alpha 4π

[1] Total efficiency is the surface efficiency times the 2π detector efficiency. The 2π alpha efficiency is based on ^{239}Pu and a surface efficiency of 0.25. The 2π beta efficiency is based on 99Tc and a surface efficiency of 0.25. 2π efficiencies are estimated from the 4π efficiencies listed by Ludlum Measurements, Inc.

Table 2-2: Example Instrument Operating Parameters and Sensitivities

Measurement Type	Detector Model	Instrument Model	Count Time	Background[1] (cpm)	MDA (dpm/ 100 cm^2)
Total Surface Activity	Ludlum 43-93	Ludlum 2360	1 min.	300 – Beta <3 – Alpha	1110 – Beta 110 – Alpha

Total Surface Activity	Ludlum 43-93	Ludlum 2360	2 min.	300 – Beta <3 – Alpha	780 – Beta 72 – Alpha
Removable Activity (Alpha/Beta)	Ludlum 43-10-1	Ludlum 3030E	2 min. 10 min. background	80 – Beta <3 - Alpha	154 – Beta 15 – Alpha

[1] Background count rates are the published typical backgrounds from the manufacturer's specifications.

The MDAs displayed in Table 2-2 are shown as examples only; illustrating the chosen instrumentation is capable of achieving MDAs at or below the release criteria. Static count times provided in Table 2-2 represent a range of values that result in MDA levels that are applicable in a specific area. Longer static count times may be required depending on the final MDA calculations. The final MDA calculations will be performed using actual background measurements and efficiency factors obtained and tested during performance of the background study.

Daily response checks are performed and documented on control charts in accordance and are used to demonstrate that radiological instrumentation is operating correctly.

Two methods are used for collecting radiation survey data (NRC, 2000):

- Direct measurements are obtained by placing the detector near or against the surface or in the media being surveyed and reading the radioactivity level directly. Count times are increased from the standard one-minute count interval as needed to achieve a MDA less than the release criteria.

- Wipes are collected to estimate the amount of removable residual radioactivity that is present on the surface.

- In addition to direct surveys and wipe samples, scan measurements may be performed at the discretion of the health physicist performing the survey. Professional judgment will be used to locate these areas based on the area being surveyed, potential historical activities, and architectural features such as cracks, seams, or other places where residual contamination could collect.

The scan survey LLD and MDC were calculated using RadPro Calculator online, with the calculations hand verified to be within the recommendations in MARSSIM. For the handheld detectors, Model 43-93 the LLD is 262 cpm with the associated MDC of 1746 dpm/100 cm^2. For the floor monitor, with a Model 43-47 detector, the LLD is 605 cpm with the associated MDC of 368 dpm/100cm^2

3.0 SURVEY INSTRUCTIONS

Direct and scan surveys and wipe samples will be performed at each survey point. Each survey point will be approximately 1 square meter (1 meter on each side). All results will be logged on a data sheet.

3.1 PREREQUISITES

1. Verify the survey instrument is within the calibration due date.
2. Perform a source check on the survey instrument. Acceptance criteria are ±20% of the source value. Perform a check of both alpha and beta. Record on the log sheet and plot the point on the graph. If the instrument response is outside the acceptance criteria, check the instrument for loose cables and holes in the window. Repeat source check. If it fails again, tag the instrument out of service and obtain another instrument.
3. Perform a 10-minute background check for both alpha and beta. Record these results on the log sheet. Alpha and beta background must be less than 75% of the DL. Background readings exceeding this level must be evaluated by the NV5/Dade Moeller Project Manager prior to proceeding with the survey.
4. Perform a source check of the Sodium Iodide (NaI) detector. Acceptance criteria are ±20% of the source value. Record on the log sheet and plot the point on the graph. If the instrument response is outside the acceptance criteria, repeat source check. If it fails again, tag the instrument out of service and obtain another instrument.
5. Perform a background count with the NaI detector and record the value.
6. Take a 10-minute background on the Ludlum 3030E. Record on the log sheet
7. Perform a source check of the Ludlum 3030E for both alpha and beta. Record and plot results.
8. Obtain sufficient wipes and wipe envelopes for the number of survey points. Label envelopes with survey point designation.

3.2 ALPHA/BETA SCAN SURVEYS

1. Scan the area at 4 inches per second with the detector held at a distance from the surface of ¼ inch for beta/gamma and 1/8 inch for alpha for the hand held detectors, or at 10 inches per second with the detector set at approximately ¼ inch for the floor monitor.
2. If elevated readings (approximately twice background) are noted, stop and mark elevated area for as an additional static measurements as directed in section 3.4.
3. Note locations of scan readings, elevated counts, and actions taken.

3.3 GAMMA SCAN SURVEYS

1. Survey each survey unit with the NaI meter.
2. Meter should be held at approximately 6 inches from the floor.
3. Use a serpentine pattern to survey the area.
4. Note any elevated readings and mark elevated area.
5. Investigate any reading greater than two times background.

3.4 DIRECT SURVEYS

1. Locate the survey points as noted in the appropriate survey unit in Attachment 1, plus any additional elevated areas identified in section 3.2 or 3.3.
2. If scan surveys were not previously performed on the square meter surrounding the measurement location, scan the square meter and select any elevated area.
3. Perform direct measurement readings for gross alpha and gross beta at each location. Take a one-minute count at each survey location. Document each reading.
4. If any reading exceeds the action level, repeat the reading. If the reading is still above the action level, flag the location and notify the PNPF Radiation Safety Officer.

3.5 WIPE SURVEYS

1. Use a dry smear and wipe a 100 cm^2 squared area using moderate pressure. 100 cm^2 is accomplished using a 2-inch filter paper wiped over an approximately 8-inch long "S."
2. Place wipe in a labeled wipe envelope.
3. Count the wipes on the Ludlum 3030E and record the results.
4. If any wipe samples exceed the action level, obtain a wipe sample adjacent to the earlier one. If the second wipe also exceeds the action level, flag the area and notify the PNPF Radiation Safety Officer.

3.6 DISCRETIONARY WALL SURVEYS

1. Use professional judgment to locate original building wall material.
2. Perform a scan survey where possible of a square meter surrounding the survey point.
3. Obtain a wipe from the material.
4. Perform a direct survey for alpha and beta.
5. Note locations of readings and samples taken.

3.7 DOCUMENTATION

1. Record actual sample locations where surveys and wipes were obtained. As noted, movement of locations due to obstructions is approved.
2. Record all direct survey results.
3. Record wipe sample results.
4. Sign and date records and send to NV5/Dade Moeller Project Manager for review.

3.8 SURVEY REPORT

The survey report will include a narrative of actions taken and results achieved. Direct survey results will be reported for each survey point and will include identification of high, low and average readings. A comparison to the release criteria will be performed.

All wipe sample results will be reported along with a comparison to the release criteria.

A summary will be performed giving the overall results of the survey and whether release criteria were met.

All records generated during the survey will be included, e.g. calibration certificates and source check results.

REFERENCES

NV5/DM, 2020. DMA-TR-106, *Radiological Historical Site Assessment, Piqua Nuclear Power Facility*. NV5/Dade Moeller, Richland, WA.

NRC. 2000. *Multi-Agency Radiation Survey and Site Investigation Manual (MARSSIM)*. NUREG-1575, Revision 1, Government Printing Office, Washington, D.C.

PNNL, 2017. *Visual Sample Plan (VSP)*. Pacific Northwest National laboratory, Richland, WA

DOE O 458.1. *Radiation Protection of the Public and the Environment*. DOE Order 458.1 Change 3, Department of Energy, January 2013

10 CFR 835 Appendix D, *Surface Contamination Values*. 74 Federal Register 18116, April 2009

ATTACHMENT 1: 56 LEVEL CLASS 1 WALL SURVEY UNITS

Random sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

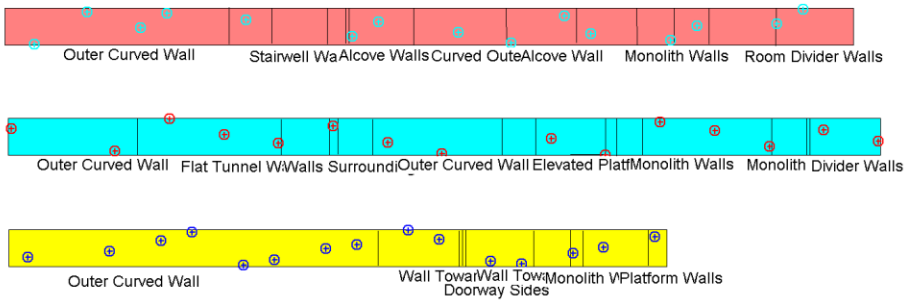
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	45
Number of selected sample areas [b]	3
Specified sampling area [c]	258.00 m^2

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.



Area: 56 Walls Class 1 #1						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
21.7008	1.9320			Random		
12.7758	0.0307			Random		
30.6258	0.6974			Random		
8.3133	1.3641			Random		
26.1633	0.2530			Random		
17.2383	0.9196			Random		
35.0883	1.5863			Random		
1.0617	0.4752			Random		
18.9117	1.1418			Random		
9.9867	1.8085			Random		
27.8367	0.1048			Random		
5.5242	0.7715			Random		
23.3742	1.4381			Random		
14.4492	0.3270			Random		
32.2992	0.9937			Random		

Area: 56 Walls Class 1 #2						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
14.6729	6.6496			Random		
38.3229	7.3162			Random		
8.7604	7.9829			Random		
32.4104	6.0158			Random		
20.5854	6.6825			Random		
44.2354	7.3492			Random		
5.8042	6.2380			Random		
29.4542	6.9047			Random		

17.6292	7.5714			Random		
41.2792	6.4603			Random		
11.7167	7.1269			Random		
35.3667	7.7936			Random		
23.5417	6.0899			Random		
47.1917	6.7566			Random		
0.1688	7.4232			Random		

Area: 56 Walls Class 1 #3						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
37.3604	13.0171			Random		
8.6104	13.6838			Random		
31.6104	12.5727			Random		
20.1104	13.2393			Random		
43.1104	13.9060			Random		
1.4229	12.0048			Random		
24.4229	12.6714			Random		
12.9229	13.3381			Random		
35.9229	12.2270			Random		
7.1729	12.8937			Random		
30.1729	13.5603			Random		
18.6729	12.4492			Random		
41.6729	13.1159			Random		
4.2979	13.7825			Random		
27.2979	12.0788			Random		

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A nonparametric random sampling approach was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, simple random point sampling was chosen. Locating the sample points randomly provides data that are separated by varying distances, providing good information about the spatial structure of the potential contamination. Knowledge of the spatial structure is useful for geostatistical analysis. However, it may not ensure that all portions of the site are equally represented.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,
β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	**n[a]**	**n[b]**	**n[c]**	**Parameter**					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]

Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42
	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that $\mu >$ action level
α = Alpha (%), Probability of mistakenly concluding that $\mu <$ action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 2: 56 LEVEL CLASS 1 SURVEY UNITS

Systematic sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

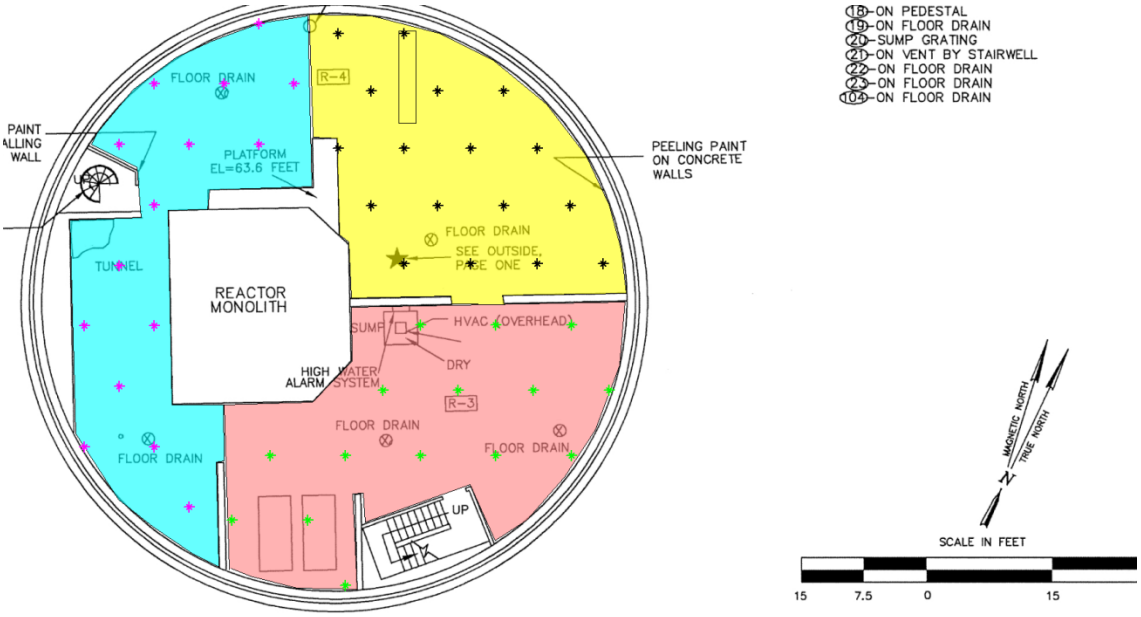
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Systematic with a random start location
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	47
Number of selected sample areas [b]	3
Specified sampling area [c]	2888.13 ft^2
Size of grid / Area of grid cell [d]	8.99072 feet / 70.0035 ft^2
Grid pattern	Triangular

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Size of grid / Area of grid gives the linear and square dimensions of the grid used to systematically place samples. If there was more than one sample area, this represents the largest dimensions used.



Area: 56 Class 1 #1						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
6.6246	4.2513			Systematic		
14.5472	4.2513			Systematic		
22.4698	4.2513			Systematic		
30.3924	4.2513			Systematic		
2.6632	11.1124			Systematic		
10.5859	11.1124			Systematic		
18.5085	11.1124			Systematic		
26.4311	11.1124			Systematic		
-1.2981	17.9736			Systematic		
6.6246	17.9736			Systematic		
14.5472	17.9736			Systematic		
22.4698	17.9736			Systematic		
2.6632	24.8348			Systematic		
10.5859	24.8348			Systematic		
18.5085	24.8348			Systematic		
-1.2981	31.6960			Systematic		
6.6246	31.6960			Systematic		

Area: 56 Class 1 #2						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
-19.0393	-24.8493			Systematic		
-31.5390	-17.6327			Systematic		
-23.2059	-17.6327			Systematic		
-27.3724	-10.4160			Systematic		
-31.5390	-3.1994			Systematic		
-23.2059	-3.1994			Systematic		

-27.3724	4.0173			Systematic		
-23.2059	11.2340			Systematic		
-27.3724	18.4506			Systematic		
-19.0393	18.4506			Systematic		
-10.7063	18.4506			Systematic		
-23.2059	25.6673			Systematic		
-14.8728	25.6673			Systematic		
-6.5397	25.6673			Systematic		
-10.7063	32.8839			Systematic		

Area: 56 Class 1 #3						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
-0.4227	-34.2147			Systematic		
-13.9088	-26.4285			Systematic		
-4.9181	-26.4285			Systematic		
-9.4134	-18.6423			Systematic		
-0.4227	-18.6423			Systematic		
8.5680	-18.6423			Systematic		
17.5587	-18.6423			Systematic		
26.5495	-18.6423			Systematic		
4.0727	-10.8561			Systematic		
13.0634	-10.8561			Systematic		
22.0541	-10.8561			Systematic		
31.0448	-10.8561			Systematic		
8.5680	-3.0699			Systematic		
17.5587	-3.0699			Systematic		
26.5495	-3.0699			Systematic		

Primary Sampling Objective

The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach

A nonparametric systematic sampling approach with a random start was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, systematic grid point sampling was chosen. Locating the sample points systematically provides data that are all equidistant apart. This approach does not provide as much information about the spatial structure of the potential contamination as simple random sampling does. Knowledge of the spatial structure is useful for geostatistical analysis. However, it

ensures that all portions of the site are equally represented. Statistical analyses of systematically collected data are valid if a random start to the grid is used.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,
β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	n^a	n^b	n^c	Parameter					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance
The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected probabilistically.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the gridded sample locations were selected based on a random start.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42

	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 3: 79 LEVEL CLASS 1 WALL SURVEY UNITS

Random sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

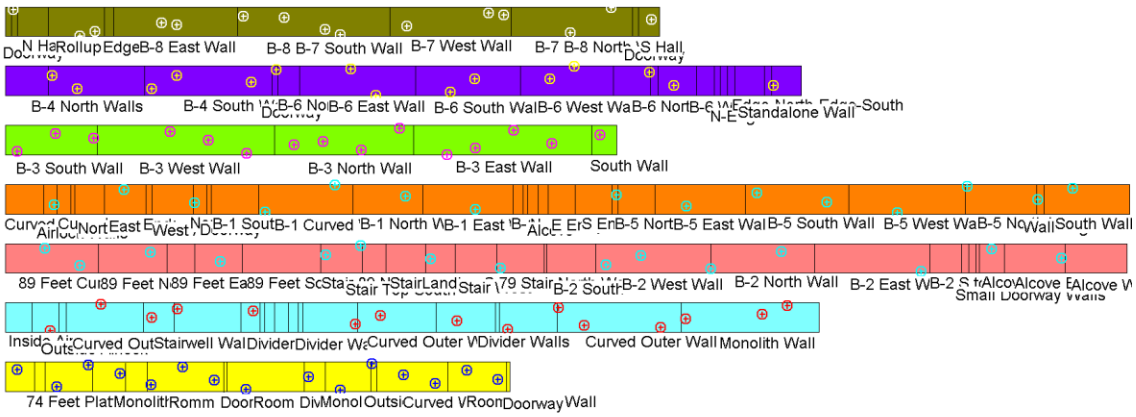
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	105
Number of selected sample areas [b]	7
Specified sampling area [c]	760.40 m^2

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A nonparametric random sampling approach was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, simple random point sampling was chosen. Locating the sample points randomly provides data that are separated by varying distances, providing good information about the spatial structure of the potential contamination. Knowledge of the spatial structure is useful for geostatistical analysis. However, it may not ensure that all portions of the site are equally represented.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,
β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	**n[a]**	**n[b]**	**n[c]**	**Parameter**					
				S_{total}	**Δ**	**α**	**β**	**$Z_{1-\alpha}$** [d]	**$Z_{1-\beta}$** [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.

[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance
The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:
1. the computed sign test statistic is normally distributed,

2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42
	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that $\mu >$ action level
α = Alpha (%), Probability of mistakenly concluding that $\mu <$ action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 4: 79 LEVEL CLASS 1 SURVEY UNITS

Systematic sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

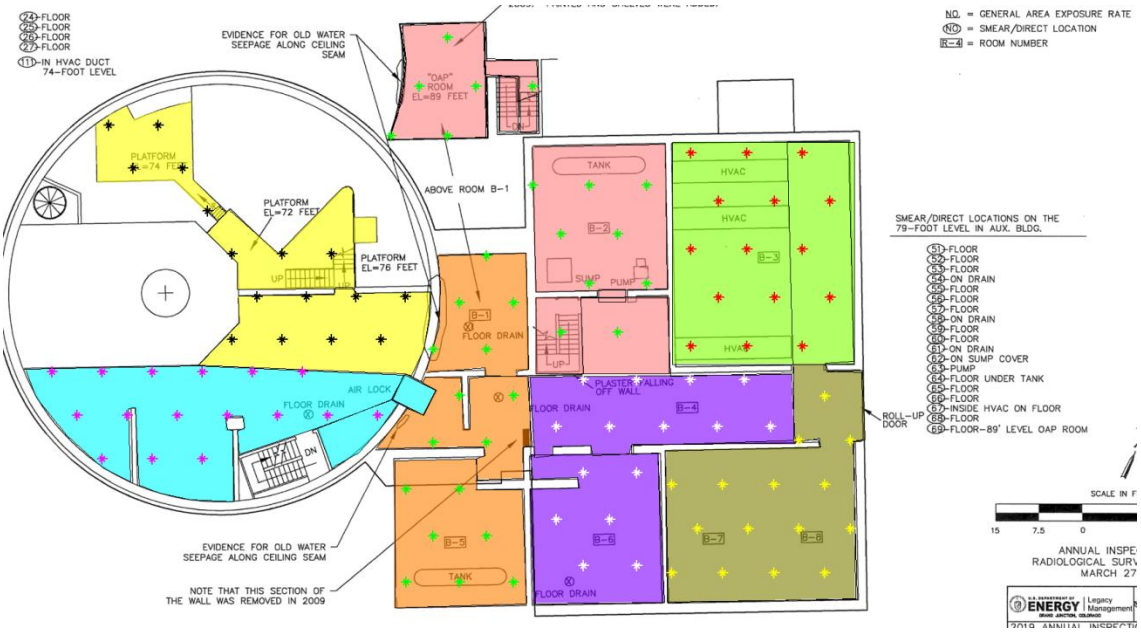
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Systematic with a random start location
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	108
Number of selected sample areas [b]	7
Specified sampling area [c]	731.63 m^2
Size of grid / Area of grid cell [d]	3.04025 meters / 8.00476 m^2
Grid pattern	Triangular

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Size of grid / Area of grid gives the linear and square dimensions of the grid used to systematically place samples. If there was more than one sample area, this represents the largest dimensions used.



Primary Sampling Objective

The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach

A nonparametric systematic sampling approach with a random start was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, systematic grid point sampling was chosen. Locating the sample points systematically provides data that are all equidistant apart. This approach does not provide as much information about the spatial structure of the potential contamination as simple random sampling does. Knowledge of the spatial structure is useful for geostatistical analysis. However, it ensures that all portions of the site are equally represented. Statistical analyses of systematically collected data are valid if a random start to the grid is used.

Nuclides

The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,
β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	**n[a]**	**n[b]**	**n[c]**	**Parameter**					
				S_{total}	**Δ**	**α**	**β**	**$Z_{1-\alpha}$** [d]	**$Z_{1-\beta}$** [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.

[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance
The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected probabilistically.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the gridded sample locations were selected based on a random start.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42
	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that $\mu >$ action level
α = Alpha (%), Probability of mistakenly concluding that $\mu <$ action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 5: 83 LEVEL CLASS 1 WALL SURVEY UNITS

Random sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

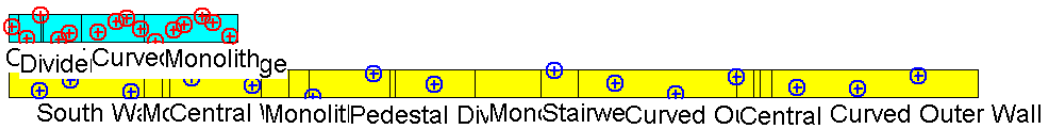
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	30
Number of selected sample areas [b]	2
Specified sampling area [c]	173.80 m^2

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.



Area: 83 Walls Class 1 #1						
X Coord	Y Coord	Label	Value	Type	Historical	Sample Area
17.6514	0.8214			Random		
52.8014	1.4881			Random		
8.8639	0.3769			Random		
44.0139	1.0436			Random		
26.4389	1.7103			Random		
61.5889	0.5992			Random		
4.4701	1.2658			Random		
39.6201	1.9325			Random		
22.0451	0.0313			Random		
57.1951	0.6979			Random		
13.2576	1.3646			Random		
48.4076	0.2535			Random		
30.8326	0.9201			Random		
65.9826	1.5868			Random		
2.2733	0.4757			Random		

Area: 83 Walls Class 1 #2						
X Coord	Y Coord	Label	Value	Type	Historical	Sample Area
14.0243	5.8831			Random		
3.6493	4.1794			Random		
11.9493	4.8461			Random		
7.7993	5.5127			Random		
16.0993	4.4016			Random		
0.2774	5.0683			Random		
8.5774	5.7350			Random		
4.4274	4.6239			Random		

12.7274	5.2905			Random		
2.3524	5.9572			Random		
10.6524	4.0560			Random		
6.5024	4.7226			Random		
14.8024	5.3893			Random		
1.3149	4.2782			Random		
9.6149	4.9448			Random		

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A nonparametric random sampling approach was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, simple random point sampling was chosen. Locating the sample points randomly provides data that are separated by varying distances, providing good information about the spatial structure of the potential contamination. Knowledge of the spatial structure is useful for geostatistical analysis. However, it may not ensure that all portions of the site are equally represented.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,
β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	n[a]	n[b]	n[c]	Parameter					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

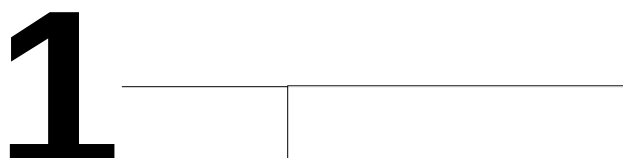
[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the

lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.



Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ > action level and alpha (%), probability of mistakenly concluding that μ < action level. The following table shows the results of this analysis.

Number of Samples			
AL=1250	**α=5**	**α=10**	**α=15**

		s=480	s=240	s=480	s=240	s=480	s=240
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42
	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 6: 83 LEVEL CLASS 1 SURVEY UNITS

Systematic sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

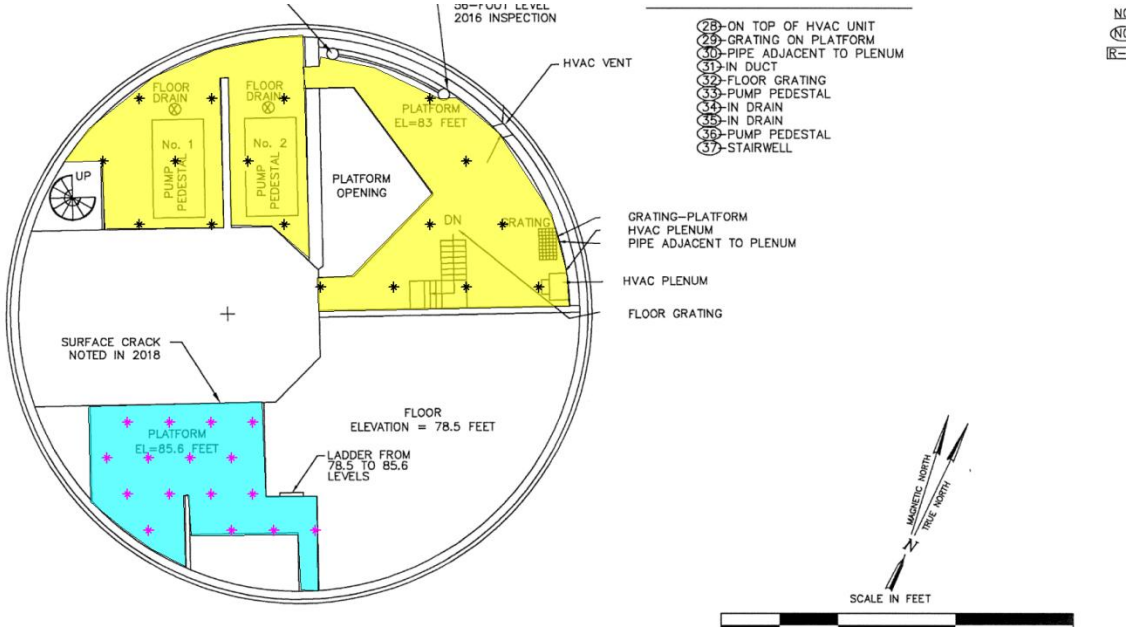
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Systematic with a random start location
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	33
Number of selected sample areas [b]	2
Specified sampling area [c]	143.07 m^2
Size of grid / Area of grid cell [d]	2.82745 meters / 6.92342 m^2
Grid pattern	Triangular

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Size of grid / Area of grid gives the linear and square dimensions of the grid used to systematically place samples. If there was more than one sample area, this represents the largest dimensions used.



Area: 83 Class 1 #1						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
0.0071	0.9023			Systematic		
2.8345	0.9023			Systematic		
5.6620	0.9023			Systematic		
8.4894	0.9023			Systematic		
-7.0616	3.3510			Systematic		
-4.2341	3.3510			Systematic		
-1.4067	3.3510			Systematic		
4.2482	3.3510			Systematic		
7.0757	3.3510			Systematic		
-8.4753	5.7996			Systematic		
-5.6478	5.7996			Systematic		
-2.8204	5.7996			Systematic		
5.6620	5.7996			Systematic		
-7.0616	8.2483			Systematic		
-4.2341	8.2483			Systematic		
-1.4067	8.2483			Systematic		
4.2482	8.2483			Systematic		

Area: 83 Class 1 #2						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
-6.6972	-8.5585			Systematic		
-3.4536	-8.5585			Systematic		
-1.8318	-8.5585			Systematic		
-0.2100	-8.5585			Systematic		
-7.5081	-7.1540			Systematic		
-5.8863	-7.1540			Systematic		

-4.2645	-7.1540			Systematic		
-2.6427	-7.1540			Systematic		
-8.3190	-5.7495			Systematic		
-6.6972	-5.7495			Systematic		
-5.0754	-5.7495			Systematic		
-3.4536	-5.7495			Systematic		
-7.5081	-4.3450			Systematic		
-5.8863	-4.3450			Systematic		
-4.2645	-4.3450			Systematic		
-2.6427	-4.3450			Systematic		

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A nonparametric systematic sampling approach with a random start was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, systematic grid point sampling was chosen. Locating the sample points systematically provides data that are all equidistant apart. This approach does not provide as much information about the spatial structure of the potential contamination as simple random sampling does. Knowledge of the spatial structure is useful for geostatistical analysis. However, it ensures that all portions of the site are equally represented. Statistical analyses of systematically collected data are valid if a random start to the grid is used.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated

so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,
β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	n[a]	n[b]	n[c]	Parameter					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance
The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected probabilistically.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the gridded sample locations were selected based on a random start.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ >

action level and alpha (%), probability of mistakenly concluding that μ < action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		α=5		α=10		α=15	
		s=480	s=240	s=480	s=240	s=480	s=240
LBGR=90	β=5	309	83	244	66	206	56
	β=10	244	66	188	51	154	42
	β=15	206	56	154	42	123	34
LBGR=80	β=5	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	β=5	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 7: 100 LEVEL CLASS 2 WALL 3+METER SURVEY UNIT

Random sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

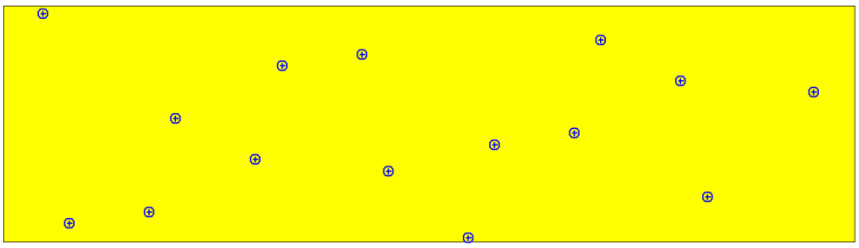
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	15
Number of selected sample areas [b]	1
Specified sampling area [c]	36000.00 ft^2

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.



Area: Area 1						
X Coord	Y Coord	Label	Value	Type	Historical	Sample Area
27.7928	7.7584			Random		
207.7928	41.0917			Random		
117.7928	74.4251			Random		
297.7928	18.8695			Random		
72.7928	52.2029			Random		
252.7928	85.5362			Random		
162.7928	29.9806			Random		
342.7928	63.3140			Random		
16.5428	96.6473			Random		
196.5428	1.5856			Random		
106.5428	34.9189			Random		
286.5428	68.2522			Random		
61.5428	12.6967			Random		
241.5428	46.0300			Random		
151.5428	79.3633			Random		

Primary Sampling Objective

The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach

A nonparametric random sampling approach was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and

historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, simple random point sampling was chosen. Locating the sample points randomly provides data that are separated by varying distances, providing good information about the spatial structure of the potential contamination. Knowledge of the spatial structure is useful for geostatistical analysis. However, it may not ensure that all portions of the site are equally represented.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,

β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,

$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,

$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	n[a]	n[b]	n[c]	Parameter					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42

	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 8: 100 LEVEL CLASS 1 WALL SURVEY UNITS

Random sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

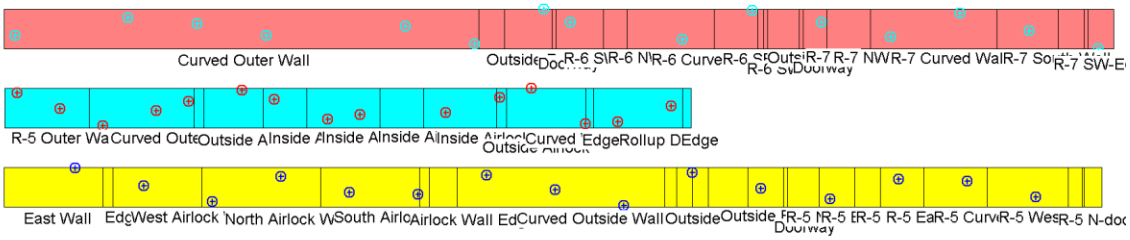
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	45
Number of selected sample areas [b]	3
Specified sampling area [c]	292.60 m^2

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.



Area: 100 Walls Class 1 #1						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
27.8770	0.8388			Random		
14.0020	1.5055			Random		
41.7520	0.3944			Random		
7.0645	1.0611			Random		
34.8145	1.7277			Random		
20.9395	0.6166			Random		
48.6895	1.2833			Random		
3.5958	1.9500			Random		
31.3458	0.0487			Random		
17.4708	0.7154			Random		
45.2208	1.3821			Random		
10.5333	0.2709			Random		
38.2833	0.9376			Random		
24.4083	1.6043			Random		
52.1583	0.4932			Random		

Area: 100 Walls Class 1 #2						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
12.0409	5.9006			Random		
29.3909	4.1969			Random		
7.7034	4.8635			Random		
25.0534	5.5302			Random		
16.3784	4.4191			Random		
33.7284	5.0858			Random		
0.6550	5.7524			Random		
18.0050	4.6413			Random		

9.3300	5.3080			Random		
26.6800	5.9746			Random		
4.9925	4.0734			Random		
22.3425	4.7401			Random		
13.6675	5.4067			Random		
31.0175	4.2956			Random		
2.8237	4.9623			Random		

Area: 100 Walls Class 1 #3						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
44.8331	8.5919			Random		
9.7706	9.2586			Random		
37.8206	9.9253			Random		
23.7956	8.2216			Random		
51.8456	8.8882			Random		
6.2644	9.5549			Random		
34.3144	8.4438			Random		
20.2894	9.1105			Random		
48.3394	9.7771			Random		
13.2769	8.6660			Random		
41.3269	9.3327			Random		
27.3019	9.9993			Random		
55.3519	8.0006			Random		
0.5667	8.6672			Random		
28.6167	9.3339			Random		

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A nonparametric random sampling approach was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, simple random point sampling was chosen. Locating the sample points randomly provides data that are separated by varying distances, providing good information about the spatial structure of the potential contamination. Knowledge of the spatial structure is useful for geostatistical analysis. However, it may not ensure that all portions of the site are equally represented.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,
β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	**n[a]**	**n[b]**	**n[c]**	**Parameter**					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]

Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42
	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that $\mu >$ action level
α = Alpha (%), Probability of mistakenly concluding that $\mu <$ action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 9: 100 LEVEL CLASS 1 SURVEY UNITS

Systematic sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

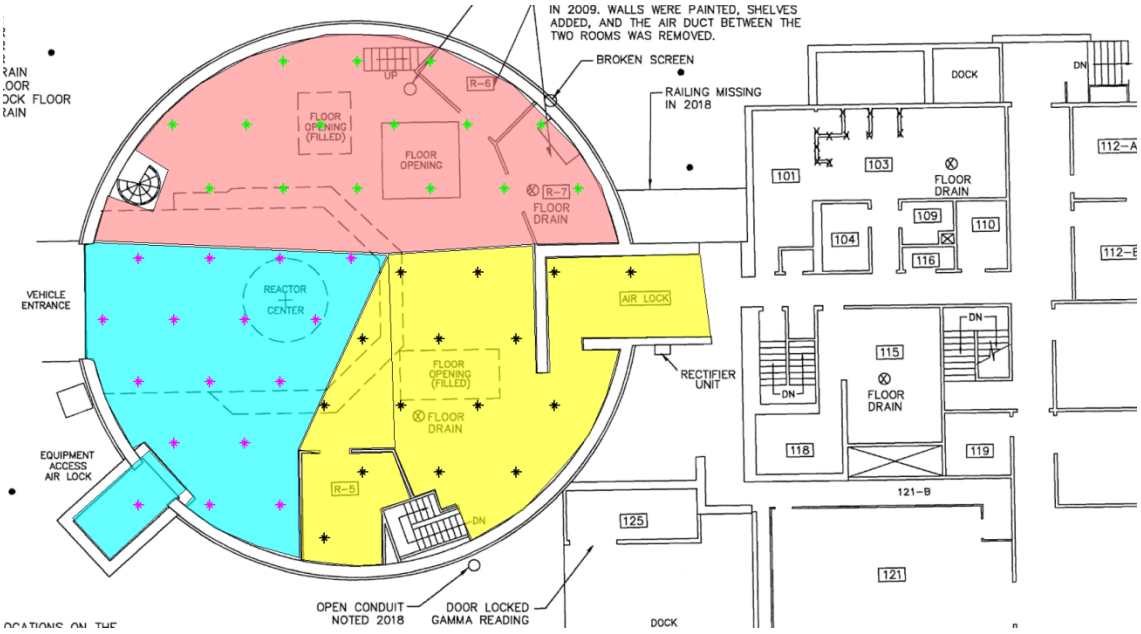
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Systematic with a random start location
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	46
Number of selected sample areas [b]	3
Specified sampling area [c]	347.45 m^2
Size of grid / Area of grid cell [d]	3.02817 meters / 7.94129 m^2
Grid pattern	Triangular

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Size of grid / Area of grid gives the linear and square dimensions of the grid used to systematically place samples. If there was more than one sample area, this represents the largest dimensions used.



Area: 100 Class 1 #1						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
9.1577	-11.6858			Systematic		
10.6718	-9.0634			Systematic		
13.7000	-9.0634			Systematic		
16.7281	-9.0634			Systematic		
9.1577	-6.4409			Systematic		
12.1859	-6.4409			Systematic		
15.2141	-6.4409			Systematic		
18.2422	-6.4409			Systematic		
10.6718	-3.8184			Systematic		
13.7000	-3.8184			Systematic		
16.7281	-3.8184			Systematic		
12.1859	-1.1960			Systematic		
15.2141	-1.1960			Systematic		
18.2422	-1.1960			Systematic		
21.2704	-1.1960			Systematic		

Area: 100 Class 1 #2						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
1.8131	-10.3604			Systematic		
4.6191	-10.3604			Systematic		
7.4251	-10.3604			Systematic		
3.2161	-7.9303			Systematic		
6.0221	-7.9303			Systematic		
1.8131	-5.5002			Systematic		
4.6191	-5.5002			Systematic		
7.4251	-5.5002			Systematic		

0.4100	-3.0701			Systematic		
3.2161	-3.0701			Systematic		
6.0221	-3.0701			Systematic		
8.8282	-3.0701			Systematic		
1.8131	-0.6400			Systematic		
4.6191	-0.6400			Systematic		
7.4251	-0.6400			Systematic		
10.2312	-0.6400			Systematic		

Area: 100 Class 1 #3						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
4.6367	2.1181			Systematic		
7.5440	2.1181			Systematic		
10.4513	2.1181			Systematic		
13.3586	2.1181			Systematic		
16.2659	2.1181			Systematic		
19.1732	2.1181			Systematic		
3.1830	4.6358			Systematic		
6.0903	4.6358			Systematic		
8.9976	4.6358			Systematic		
11.9049	4.6358			Systematic		
14.8122	4.6358			Systematic		
17.7195	4.6358			Systematic		
7.5440	7.1536			Systematic		
10.4513	7.1536			Systematic		
13.3586	7.1536			Systematic		

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A nonparametric systematic sampling approach with a random start was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, systematic grid point sampling was chosen. Locating the sample points systematically provides data that are all equidistant apart. This approach does not provide as much information about the spatial structure of the potential contamination as simple random sampling does. Knowledge of the spatial structure is useful for geostatistical analysis. However, it

ensures that all portions of the site are equally represented. Statistical analyses of systematically collected data are valid if a random start to the grid is used.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,
β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	n^a	n^b	n^c	Parameter					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance
The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected probabilistically.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the gridded sample locations were selected based on a random start.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42

	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 10: 100 LEVEL CLASS 2 AND 3 SURVEY UNITS

Systematic sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

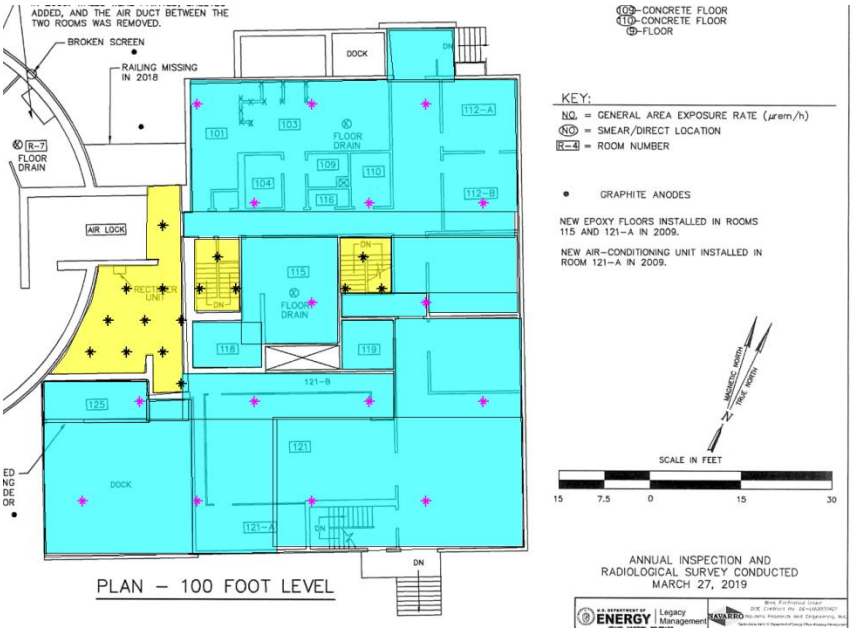
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Systematic with a random start location
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	32
Number of selected sample areas [b]	2
Specified sampling area [c]	478.03 m^2
Size of grid / Area of grid cell [d]	5.74203 meters / 28.5536 m^2
Grid pattern	Triangular

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Size of grid / Area of grid gives the linear and square dimensions of the grid used to systematically place samples. If there was more than one sample area, this represents the largest dimensions used.



Area: 100 Class 2						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
-1613.7050	-1250.4414			Systematic		
-1618.2700	-1248.8600			Systematic		
-1616.4440	-1248.8600			Systematic		
-1614.6180	-1248.8600			Systematic		
-1617.3570	-1247.2786			Systematic		
-1615.5310	-1247.2786			Systematic		
-1613.7050	-1247.2786			Systematic		
-1616.4440	-1245.6973			Systematic		
-1614.6180	-1245.6973			Systematic		
-1612.7920	-1245.6973			Systematic		
-1610.9660	-1245.6973			Systematic		
-1605.4881	-1245.6973			Systematic		
-1603.6621	-1245.6973			Systematic		
-1611.8790	-1244.1159			Systematic		
-1604.5751	-1244.1159			Systematic		
-1614.6180	-1242.5346			Systematic		

Area: 100 Class 3						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
-1618.6571	-1256.3286			Systematic		
-1612.9150	-1256.3286			Systematic		
-1607.1730	-1256.3286			Systematic		
-1601.4310	-1256.3286			Systematic		
-1615.7861	-1251.3559			Systematic		
-1610.0440	-1251.3559			Systematic		
-1604.3020	-1251.3559			Systematic		

-1598.5600	-1251.3559			Systematic		
-1607.1730	-1246.3831			Systematic		
-1601.4310	-1246.3831			Systematic		
-1610.0440	-1241.4104			Systematic		
-1604.3020	-1241.4104			Systematic		
-1598.5600	-1241.4104			Systematic		
-1612.9150	-1236.4376			Systematic		
-1607.1730	-1236.4376			Systematic		
-1601.4310	-1236.4376			Systematic		

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A nonparametric systematic sampling approach with a random start was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, systematic grid point sampling was chosen. Locating the sample points systematically provides data that are all equidistant apart. This approach does not provide as much information about the spatial structure of the potential contamination as simple random sampling does. Knowledge of the spatial structure is useful for geostatistical analysis. However, it ensures that all portions of the site are equally represented. Statistical analyses of systematically collected data are valid if a random start to the grid is used.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated

so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,
β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	n[a]	n[b]	n[c]	Parameter					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance
The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected probabilistically.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the gridded sample locations were selected based on a random start.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ >

action level and alpha (%), probability of mistakenly concluding that μ < action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		α=5		α=10		α=15	
		s=480	s=240	s=480	s=240	s=480	s=240
LBGR=90	β=5	309	83	244	66	206	56
	β=10	244	66	188	51	154	42
	β=15	206	56	154	42	123	34
LBGR=80	β=5	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	β=5	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 11: 100 LEVEL CLASS 2 WALL SURVEY UNIT

Random sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

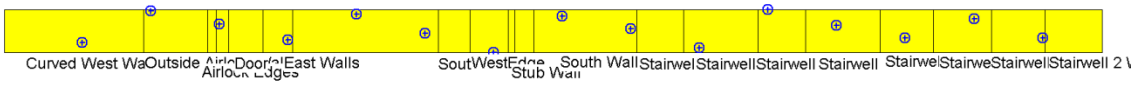
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	15
Number of selected sample areas [b]	1
Specified sampling area [c]	102.80 m^2

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.



Area: 100 Walls Class 2						
X Coord	Y Coord	Label	Value	Type	Historical	Sample Area
26.1228	1.7031			Random		
13.2728	0.5920			Random		
38.9728	1.2587			Random		
6.8478	1.9253			Random		
32.5478	0.2216			Random		
19.6978	0.8883			Random		
45.3978	1.5550			Random		
3.6353	0.4439			Random		
29.3353	1.1105			Random		
16.4853	1.7772			Random		
42.1853	0.6661			Random		
10.0603	1.3328			Random		
35.7603	1.9994			Random		
22.9103	0.0006			Random		
48.6103	0.6673			Random		

Primary Sampling Objective

The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach

A nonparametric random sampling approach was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and

historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, simple random point sampling was chosen. Locating the sample points randomly provides data that are separated by varying distances, providing good information about the spatial structure of the potential contamination. Knowledge of the spatial structure is useful for geostatistical analysis. However, it may not ensure that all portions of the site are equally represented.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,

β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,

$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,

$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	n[a]	n[b]	n[c]	Parameter					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42

	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 12: 111 LEVEL CLASS 1 AND 3 SURVEY UNITS

Systematic sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

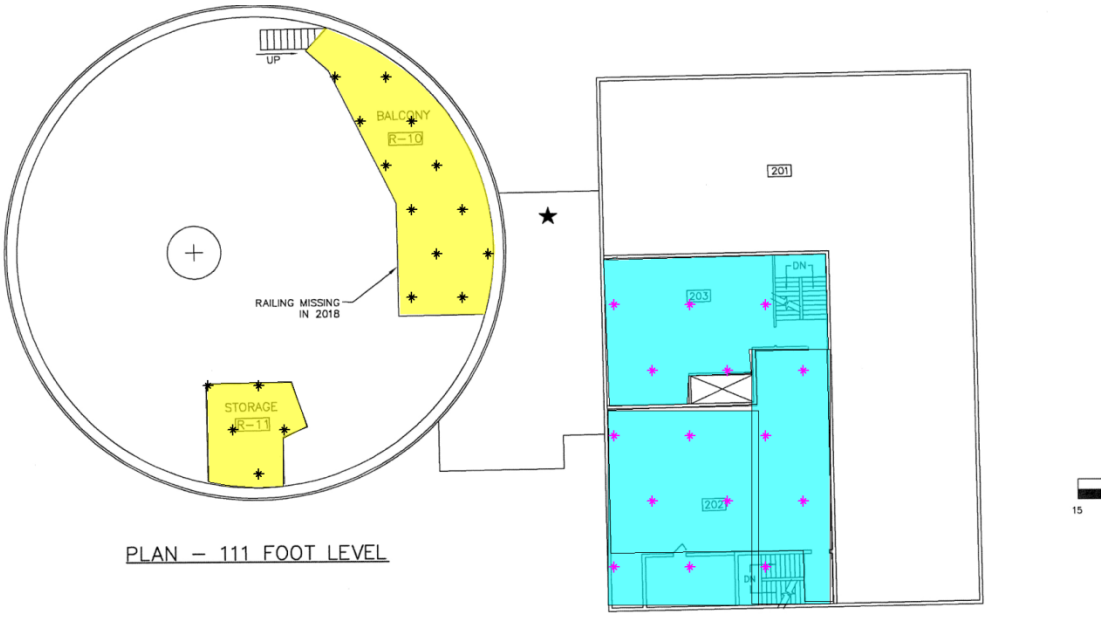
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Systematic with a random start location
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	32
Number of selected sample areas [b]	2
Specified sampling area [c]	215.84 m^2
Size of grid / Area of grid cell [d]	3.38133 meters / 9.90162 m^2
Grid pattern	Triangular

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Size of grid / Area of grid gives the linear and square dimensions of the grid used to systematically place samples. If there was more than one sample area, this represents the largest dimensions used.



Area: 111 Class 1						
X Coord	Y Coord	Label	Value	Type	Historical	Sample Area
-1632.7482	-1256.7238			Systematic		
-1633.8864	-1254.7524			Systematic		
-1631.6100	-1254.7524			Systematic		
-1635.0246	-1252.7810			Systematic		
-1632.7482	-1252.7810			Systematic		
-1625.9189	-1248.8381			Systematic		
-1623.6425	-1248.8381			Systematic		
-1624.7807	-1246.8666			Systematic		
-1622.5042	-1246.8666			Systematic		
-1625.9189	-1244.8952			Systematic		
-1623.6425	-1244.8952			Systematic		
-1627.0571	-1242.9237			Systematic		
-1624.7807	-1242.9237			Systematic		
-1628.1953	-1240.9523			Systematic		
-1625.9189	-1240.9523			Systematic		
-1629.3335	-1238.9808			Systematic		
-1627.0571	-1238.9808			Systematic		

Area: 111 Class 3						
X Coord	Y Coord	Label	Value	Type	Historical	Sample Area
-1616.8491	-1260.8767			Systematic		
-1613.4678	-1260.8767			Systematic		
-1610.0865	-1260.8767			Systematic		
-1615.1585	-1257.9484			Systematic		
-1611.7771	-1257.9484			Systematic		
-1608.3958	-1257.9484			Systematic		

-1616.8491	-1255.0201			Systematic		
-1613.4678	-1255.0201			Systematic		
-1610.0865	-1255.0201			Systematic		
-1615.1585	-1252.0917			Systematic		
-1611.7771	-1252.0917			Systematic		
-1608.3958	-1252.0917			Systematic		
-1616.8491	-1249.1634			Systematic		
-1613.4678	-1249.1634			Systematic		
-1610.0865	-1249.1634			Systematic		

Primary Sampling Objective
The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A nonparametric systematic sampling approach with a random start was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, systematic grid point sampling was chosen. Locating the sample points systematically provides data that are all equidistant apart. This approach does not provide as much information about the spatial structure of the potential contamination as simple random sampling does. Knowledge of the spatial structure is useful for geostatistical analysis. However, it ensures that all portions of the site are equally represented. Statistical analyses of systematically collected data are valid if a random start to the grid is used.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated

so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,
β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	n[a]	n[b]	n[c]	Parameter					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance
The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected probabilistically.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the gridded sample locations were selected based on a random start.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that μ >

action level and alpha (%), probability of mistakenly concluding that μ < action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		α=5		α=10		α=15	
		s=480	s=240	s=480	s=240	s=480	s=240
LBGR=90	β=5	309	83	244	66	206	56
	β=10	244	66	188	51	154	42
	β=15	206	56	154	42	123	34
LBGR=80	β=5	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	β=5	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 13: 111 LEVEL CLASS 1 WALL SURVEY UNIT

Random sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

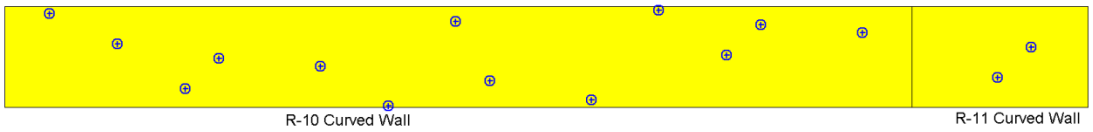
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	15
Number of selected sample areas [b]	1
Specified sampling area [c]	43.00 m^2

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.



Area: 111 Wall Class 1 #1						
X Coord	Y Coord	Label	Value	Type	Historical	Sample Area
4.2484	0.9635			Random		
14.9984	1.6302			Random		
9.6234	0.5191			Random		
20.3734	1.1857			Random		
0.8891	1.8524			Random		
11.6391	0.1487			Random		
6.2641	0.8154			Random		
17.0141	1.4820			Random		
3.5766	0.3709			Random		
14.3266	1.0376			Random		
8.9516	1.7043			Random		
19.7016	0.5932			Random		
2.2328	1.2598			Random		
12.9828	1.9265			Random		
7.6078	0.0252			Random		

Primary Sampling Objective

The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach

A nonparametric random sampling approach was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and

historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, simple random point sampling was chosen. Locating the sample points randomly provides data that are separated by varying distances, providing good information about the spatial structure of the potential contamination. Knowledge of the spatial structure is useful for geostatistical analysis. However, it may not ensure that all portions of the site are equally represented.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,

β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,

$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,

$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	n[a]	n[b]	n[c]	Parameter					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42

	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 14: 121 LEVEL CLASS 3 SURVEY UNIT

Random sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

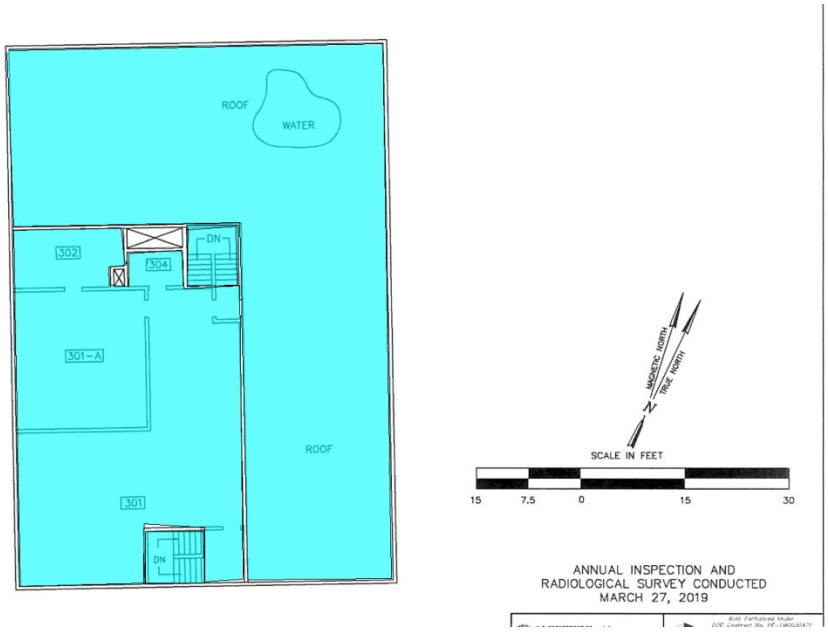
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field is also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Systematic with a random start location
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	0
Number of selected sample areas [b]	1
Specified sampling area [c]	374.21 m^2

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.



Primary Sampling Objective
The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach
A nonparametric systematic sampling approach with a random start was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, systematic grid point sampling was chosen. Locating the sample points systematically provides data that are all equidistant apart. This approach does not provide as much information about the spatial structure of the potential contamination as simple random sampling does. Knowledge of the spatial structure is useful for geostatistical analysis. However, it ensures that all portions of the site are equally represented. Statistical analyses of systematically collected data are valid if a random start to the grid is used.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,
β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,
$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,
$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	**n[a]**	**n[b]**	**n[c]**	**Parameter**					
				S_{total}	**Δ**	**α**	**β**	**$Z_{1-\alpha}$** [d]	**$Z_{1-\beta}$** [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.

[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance
The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at 1-α on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at 1-α. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions
The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected probabilistically.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the gridded sample locations were selected based on a random start.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42
	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that $\mu >$ action level
α = Alpha (%), Probability of mistakenly concluding that $\mu <$ action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 15: GROUND, DOCKS, AND ENTRIES OUTSIDE

Systematic sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

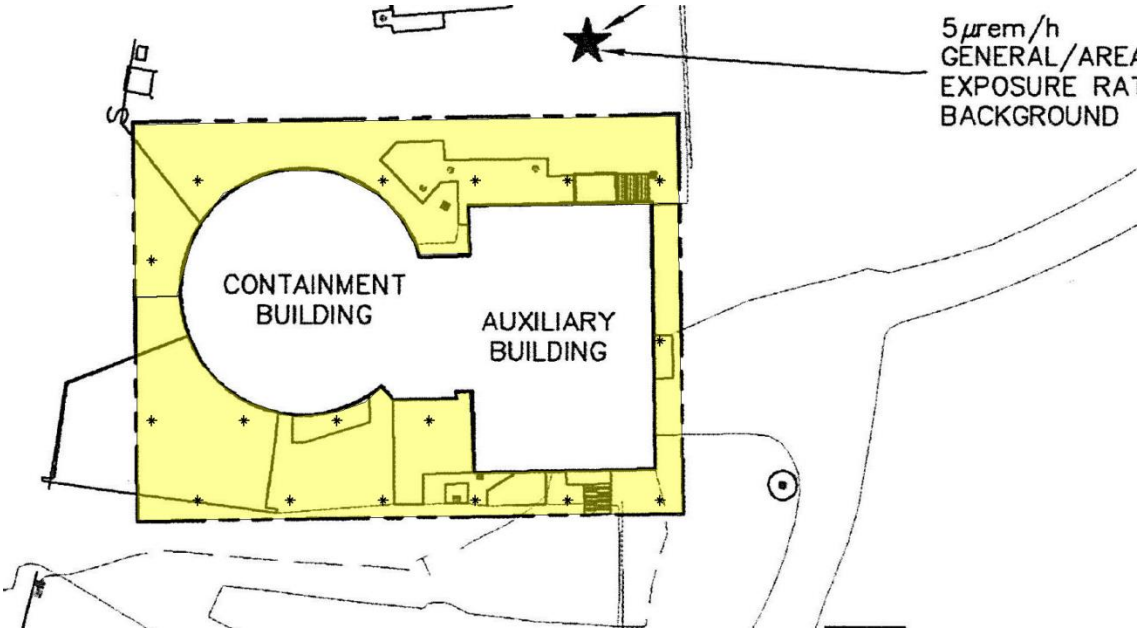
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Systematic with a random start location
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	17
Number of selected sample areas [b]	1
Specified sampling area [c]	959.96 m^2
Size of grid / Area of grid cell [d]	8.5964 meters / 63.9976 m^2
Grid pattern	Triangular

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.
[d] Size of grid / Area of grid gives the linear and square dimensions of the grid used to systematically place samples. If there was more than one sample area, this represents the largest dimensions used.



Area: Ground-Dock-Entry Class 3						
X Coord	**Y Coord**	**Label**	**Value**	**Type**	**Historical**	**Sample Area**
5.4703	1.9352			Systematic		
14.0667	1.9352			Systematic		
22.6631	1.9352			Systematic		
31.2595	1.9352			Systematic		
39.8559	1.9352			Systematic		
48.4523	1.9352			Systematic		
1.1721	9.3799			Systematic		
9.7685	9.3799			Systematic		
18.3649	9.3799			Systematic		
26.9613	9.3799			Systematic		
48.4523	16.8246			Systematic		
1.1721	24.2693			Systematic		
5.4703	31.7140			Systematic		
22.6631	31.7140			Systematic		
31.2595	31.7140			Systematic		
39.8559	31.7140			Systematic		
48.4523	31.7140			Systematic		

Primary Sampling Objective

The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach

A nonparametric systematic sampling approach with a random start was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, systematic grid point sampling was chosen. Locating the sample points systematically provides data that are all equidistant apart. This approach does not provide as much information about the spatial structure of the potential contamination as simple random sampling does. Knowledge of the spatial structure is useful for geostatistical analysis. However, it ensures that all portions of the site are equally represented. Statistical analyses of systematically collected data are valid if a random start to the grid is used.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{\left(Z_{1-\alpha} + Z_{1-\beta}\right)^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,

α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,

β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,

$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,

$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	n[a]	n[b]	n[c]	Parameter					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected probabilistically.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the gridded sample locations were selected based on a random start.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42

	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 16: AUXILIARY BUILDING OUTSIDE

Random sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

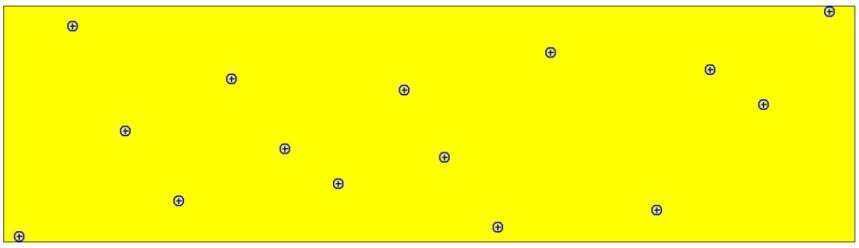
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	15
Number of selected sample areas [b]	1
Specified sampling area [c]	36000.00 ft^2

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.



Area: Area 1						
X Coord	Y Coord	Label	Value	Type	Historical	Sample Area
169.6739	64.1232			Random		
349.6739	97.4565			Random		
6.5489	2.3948			Random		
186.5489	35.7281			Random		
96.5489	69.0615			Random		
276.5489	13.5059			Random		
51.5489	46.8392			Random		
231.5489	80.1726			Random		
141.5489	24.6170			Random		
321.5489	57.9503			Random		
29.0489	91.2837			Random		
209.0489	6.0985			Random		
119.0489	39.4318			Random		
299.0489	72.7652			Random		
74.0489	17.2096			Random		

Primary Sampling Objective

The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach

A nonparametric random sampling approach was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and

historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, simple random point sampling was chosen. Locating the sample points randomly provides data that are separated by varying distances, providing good information about the spatial structure of the potential contamination. Knowledge of the spatial structure is useful for geostatistical analysis. However, it may not ensure that all portions of the site are equally represented.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,

β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,

$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,

$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	n[a]	n[b]	n[c]	Parameter					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42

	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

ATTACHMENT 17: REACTOR BUILDING OUTSIDE

Random sampling locations for comparing a median with a fixed threshold (nonparametric - MARSSIM)

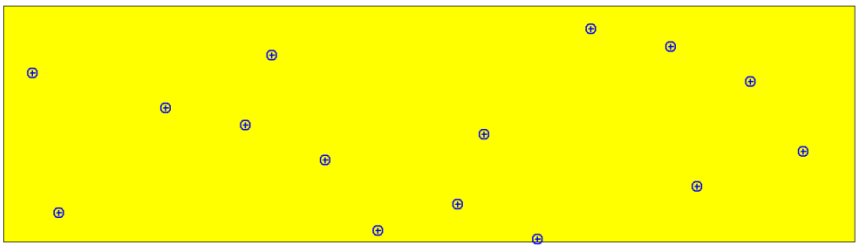
Summary

This report summarizes the sampling design used, associated statistical assumptions, as well as general guidelines for conducting post-sampling data analysis. Sampling plan components presented here include how many sampling locations to choose and where within the sampling area to collect those samples. The type of medium to sample (i.e., soil, groundwater, etc.) and how to analyze the samples (in-situ, fixed laboratory, etc.) are addressed in other sections of the sampling plan.

The following table summarizes the sampling design developed. A figure that shows sampling locations in the field and a table that lists sampling location coordinates are also provided below.

SUMMARY OF SAMPLING DESIGN	
Primary Objective of Design	Compare a site mean or median to a fixed threshold
Type of Sampling Design	Nonparametric
Sample Placement (Location) in the Field	Simple random sampling
Working (Null) Hypothesis	The median(mean) value at the site exceeds the threshold
Formula for calculating number of sampling locations	Sign Test - MARSSIM version
Calculated number of samples	12
Number of samples adjusted for EMC	12
Number of samples with MARSSIM Overage	15
Number of samples on map [a]	15
Number of selected sample areas [b]	1
Specified sampling area [c]	36000.00 ft^2

[a] This number may differ from the calculated number because of 1) grid edge effects, 2) adding judgment samples, or 3) selecting or unselecting sample areas.
[b] The number of selected sample areas is the number of colored areas on the map of the site. These sample areas contain the locations where samples are collected.
[c] The sampling area is the total surface area of the selected colored sample areas on the map of the site.



Area: Area 1						
X Coord	Y Coord	Label	Value	Type	Historical	Sample Area
225.9239	1.1602			Random		
135.9239	34.4935			Random		
315.9239	67.8269			Random		
23.4239	12.2713			Random		
203.4239	45.6047			Random		
113.4239	78.9380			Random		
293.4239	23.3824			Random		
68.4239	56.7158			Random		
248.4239	90.0491			Random		
158.4239	4.8639			Random		
338.4239	38.1973			Random		
12.1739	71.5306			Random		
192.1739	15.9750			Random		
102.1739	49.3084			Random		
282.1739	82.6417			Random		

Primary Sampling Objective

The primary purpose of sampling at this site is to compare a site median or mean value with a fixed threshold. The working hypothesis (or 'null' hypothesis) is that the median(mean) value at the site is equal to or exceeds the threshold. The alternative hypothesis is that the median(mean) value is less than the threshold. VSP calculates the number of samples required to reject the null hypothesis in favor of the alternative one, given a selected sampling approach and inputs to the associated equation.

Selected Sampling Approach

A nonparametric random sampling approach was used to determine the number of samples and to specify sampling locations. A nonparametric formula was chosen because the conceptual model and

historical information (e.g., historical data from this site or a very similar site) indicate that typical parametric assumptions may not be true.

Both parametric and non-parametric equations rely on assumptions about the population. Typically, however, non-parametric equations require fewer assumptions and allow for more uncertainty about the statistical distribution of values at the site. The trade-off is that if the parametric assumptions are valid, the required number of samples is usually less than if a non-parametric equation was used.

VSP offers many options to determine the locations at which measurements are made or samples are collected and subsequently measured. For this design, simple random point sampling was chosen. Locating the sample points randomly provides data that are separated by varying distances, providing good information about the spatial structure of the potential contamination. Knowledge of the spatial structure is useful for geostatistical analysis. However, it may not ensure that all portions of the site are equally represented.

Nuclides
The following table summarizes the analyzed nuclides.

Nuclides Analyzed by Study		
Nuclide	**$DCGL_W$**	**$DCGL_{EMC}$**
Analyte 1	1250	

Number of Total Samples: Calculation Equation and Inputs
The equation used to calculate the number of samples is based on a Sign test (see PNNL 13450 for discussion). For this site, the null hypothesis is rejected in favor of the alternative one if the median(mean) is sufficiently smaller than the threshold. The number of samples to collect is calculated so that if the inputs to the equation are true, the calculated number of samples will cause the null hypothesis to be rejected.

The formula used to calculate the number of samples is:

$$n = \frac{(Z_{1-\alpha} + Z_{1-\beta})^2}{4(SignP - 0.5)^2}$$

where

$$SignP = \Phi\left(\frac{\Delta}{S_{total}}\right)$$

$\Phi(z)$ is the cumulative standard normal distribution on $(-\infty,z)$ (see PNNL-13450 for details),
n is the number of samples,
S_{total} is the estimated standard deviation of the measured values including analytical error,
Δ is the width of the gray region,
α is the acceptable probability of incorrectly concluding the site median(mean) is less than the threshold,

β is the acceptable probability of incorrectly concluding the site median(mean) exceeds the threshold,

$Z_{1-\alpha}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\alpha}$ is $1-\alpha$,

$Z_{1-\beta}$ is the value of the standard normal distribution such that the proportion of the distribution less than $Z_{1-\beta}$ is $1-\beta$.

Note: MARSSIM suggests that the number of samples should be increased by at least 20% to account for missing or unusable data and uncertainty in the calculated value of n. VSP allows a user-supplied percent overage as discussed in MARSSIM (EPA 2000, p. 5-33).

For each nuclide in the **Nuclides Analyzed by Study** table, the values of these inputs that result in the calculated number of sampling locations are:

Nuclide	n[a]	n[b]	n[c]	Parameter					
				S_{total}	Δ	α	β	$Z_{1-\alpha}$ [d]	$Z_{1-\beta}$ [e]
Analyte 1	12	12	15	240	530	0.05	0.05	1.64485	1.64485

[a] The number of samples calculated by the formula.
[b] The number of samples increased by EMC calculations.
[c] The final number of samples increased by the MARSSIM Overage of 20%.
[d] This value is automatically calculated by VSP based upon the user defined value of α.
[e] This value is automatically calculated by VSP based upon the user defined value of β.

Performance

The following figure is a performance goal diagram, described in EPA's QA/G-4 guidance (EPA, 2000). It shows the probability of concluding the sample area is dirty on the vertical axis versus a range of possible true median(mean) values for the site on the horizontal axis. This graph contains all of the inputs to the number of samples equation and pictorially represents the calculation.

The red vertical line is shown at the threshold (action limit) on the horizontal axis. The width of the gray shaded area is equal to Δ; the upper horizontal dashed blue line is positioned at $1-\alpha$ on the vertical axis; the lower horizontal dashed blue line is positioned at β on the vertical axis. The vertical green line is positioned at one standard deviation below the threshold. The shape of the red curve corresponds to the estimates of variability. The calculated number of samples results in the curve that passes through the lower bound of Δ at β and the upper bound of Δ at $1-\alpha$. If any of the inputs change, the number of samples that result in the correct curve changes.

1

Statistical Assumptions

The assumptions associated with the formulas for computing the number of samples are:

1. the computed sign test statistic is normally distributed,
2. the variance estimate, S^2, is reasonable and representative of the population being sampled,
3. the population values are not spatially or temporally correlated, and
4. the sampling locations will be selected randomly.

The first three assumptions will be assessed in a post data collection analysis. The last assumption is valid because the sample locations were selected using a random process.

Sensitivity Analysis

The sensitivity of the calculation of number of samples was explored by varying the standard deviation, lower bound of gray region (% of action level), beta (%), probability of mistakenly concluding that $\mu >$ action level and alpha (%), probability of mistakenly concluding that $\mu <$ action level. The following table shows the results of this analysis.

Number of Samples							
AL=1250		**α=5**		**α=10**		**α=15**	
		s=480	**s=240**	**s=480**	**s=240**	**s=480**	**s=240**
LBGR=90	**β=5**	309	83	244	66	206	56
	β=10	244	66	188	51	154	42

	β=15	206	56	154	42	123	34
LBGR=80	**β=5**	83	27	66	22	56	18
	β=10	66	22	51	17	42	14
	β=15	56	18	42	14	34	11
LBGR=70	**β=5**	41	17	33	15	28	12
	β=10	33	15	26	11	21	9
	β=15	28	12	21	9	17	8

s = Standard Deviation
LBGR = Lower Bound of Gray Region (% of Action Level)
β = Beta (%), Probability of mistakenly concluding that μ > action level
α = Alpha (%), Probability of mistakenly concluding that μ < action level
AL = Action Level (Threshold)

Note: Values in table are not adjusted for EMC.

Delivering Solutions
Improving Lives