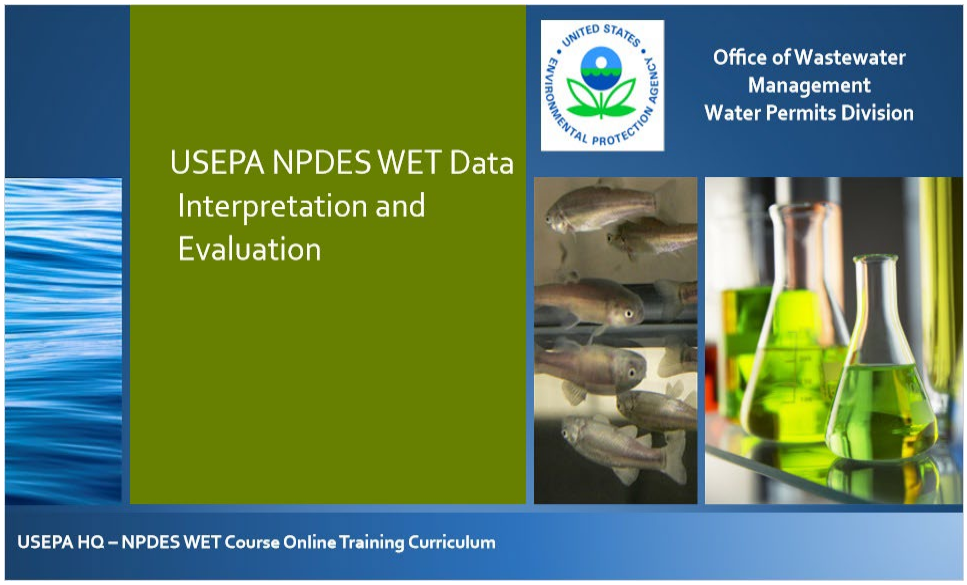
Module 5 - USEPA NPDES WET Data Interpretation and Evaluation

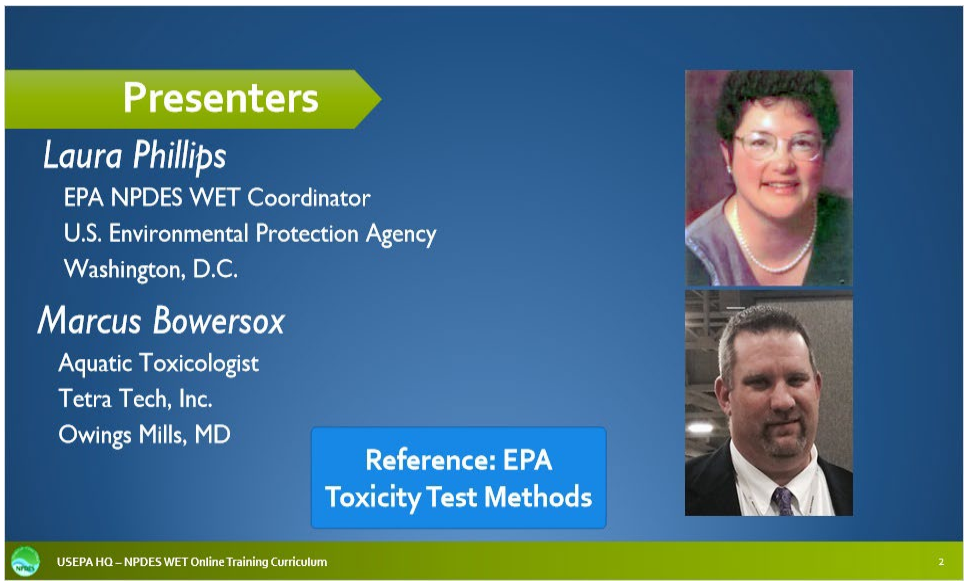


Notes:

Welcome to this presentation on the United States Environmental Protection Agency's, hereafter EPA, National Pollutant Discharge Elimination System, or NPDES, Whole Effluent Toxicity Data Interpretation and Evaluation. This presentation is part of a web-based training series on Whole Effluent Toxicity, or WET, sponsored by EPA's Office of Wastewater Management's Water Permits Division.

You can review this stand-alone presentation, or, if you have not already done so, you might also be interested in viewing the other presentations in the series, which cover the use of WET in the NPDES permit program.

Before we get started with this presentation, I will make some introductions and cover two important housekeeping items.



Notes:

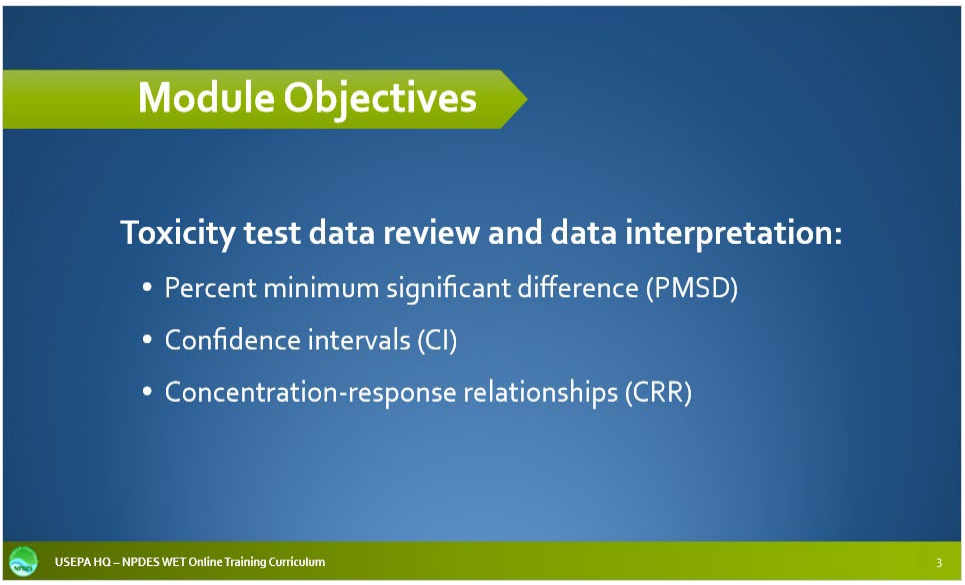
First, the introductions.
Your speakers for this presentation are, me, Laura Phillips, and I am the EPA's NPDES WET Coordinator with the Water Permits Division within the Office of Wastewater Management at EPA Headquarters in Washington, D.C., and Marcus Bowersox, EPA Headquarters contractor and an aquatic toxicologist with Tetra Tech, Incorporated in Owings Mills, Maryland. Second, now for those housekeeping items.
You should be aware that all the materials used in this presentation have been reviewed by EPA staff for technical and programmatic accuracy; however, the views of the speakers are their own and do not necessarily reflect those of EPA. The NPDES permit program, which includes the use of toxicity testing, is governed by the existing requirements of the Clean Water Act and EPA's NPDES permit implementation regulations. These statutory and regulatory provisions contain legally binding requirements. However, the information in this presentation is not binding. Furthermore, it supplements, and does not modify, existing EPA policy and guidance on WET in the NPDES permit program. EPA may revise and/or update the contents of this presentation in the future.
Throughout this module, the term "state" means a state, the District of Columbia, the territories including the Commonwealth of Puerto Rico, the United States Virgin Islands, Guam, American Samoa, the Commonwealth of the Northern Mariana Islands, and the Trust Territory of the Pacific Islands and Tribes (40 CFR Part 122.2). The term "authorized Tribe" means those federally recognized Indian Tribes with

authority to administer a Clean Water Act water quality standards, WQS, program. In some instance we may use the term “permitting authority” to include EPA, states, territories, and Tribes that have been authorized to administer the NPDES permit program.

This module was developed based on the live EPA Headquarters’ NPDES WET course that the Water Permits Division of the Office of Wastewater Management has been teaching to EPA regions, states, territories, and authorized Tribes. This course, where possible, has been developed with both the non-scientist and scientist in mind. Also, while not necessary, a basic knowledge of biological principles and WET will be helpful to the viewer. Prior to this course, a review of EPA's NPDES Permit Writers’ online course, which is available at EPA's NPDES website, is recommended. See the “Resources” tab for a link to the NPDES training website.

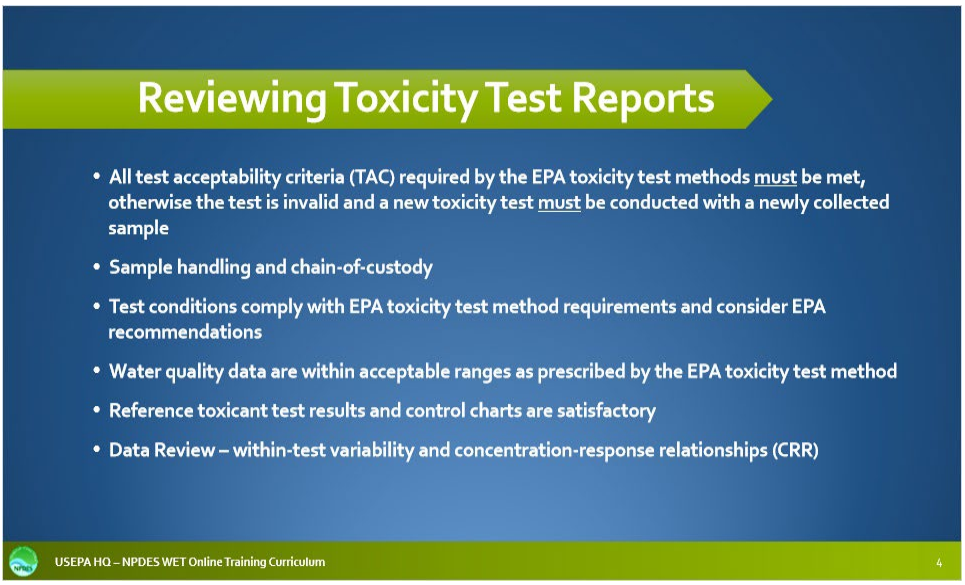
When appropriate a blue button will appear on a slide to provide access to more information. By clicking this button, additional slides will present information regarding either freshwater or marine EPA toxicity test methods. When these additional slides are finished, you will be automatically returned to the module slide where you left off. The blue button on this slide provides the references for EPA’s toxicity test methods that will be presented throughout this module.

Let me turn this over to Marcus and we will look at EPA’s NPDES WET Data Interpretation and Evaluation.



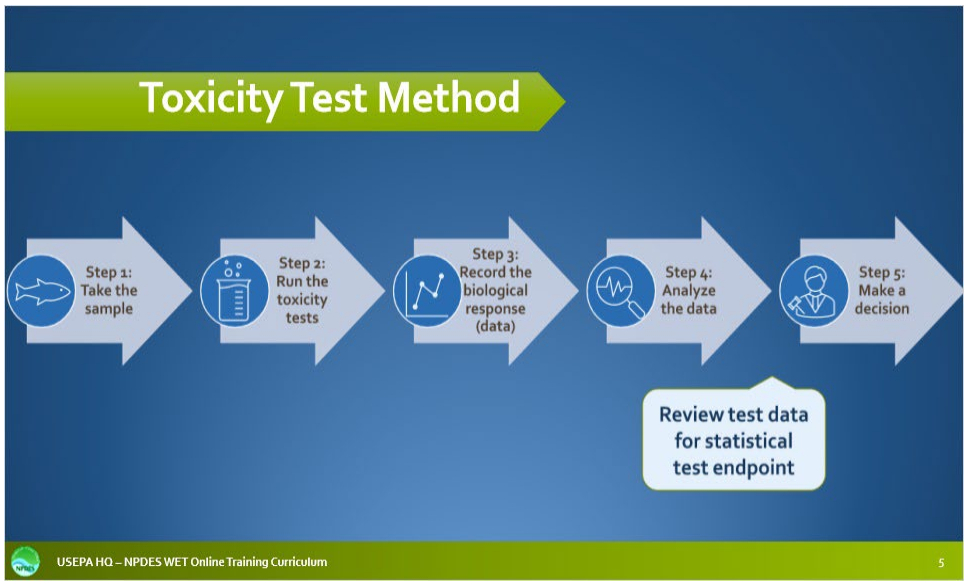
Notes:

Thanks Laura. The overall objective of this module is to review the recommended data evaluation and interpretation steps provided in the EPA toxicity test methods and in EPA's 2002 Guidelines Establishing Test Procedures for the Analysis of Pollutants; Whole Effluent Toxicity Test Methods; Final Rule, hereafter, the 2002 WET Test Methods Final Rule. These recommended data evaluation and interpretation steps include review of within-test variability, which is evaluated using the percent minimum significant difference, or PMSD, the use of confidence intervals, or CIs, with point estimate data evaluations, and as a part of the evaluation of toxicity test concentration-response relationships, or CRRs, used to help determine the reliability of valid toxicity test data.



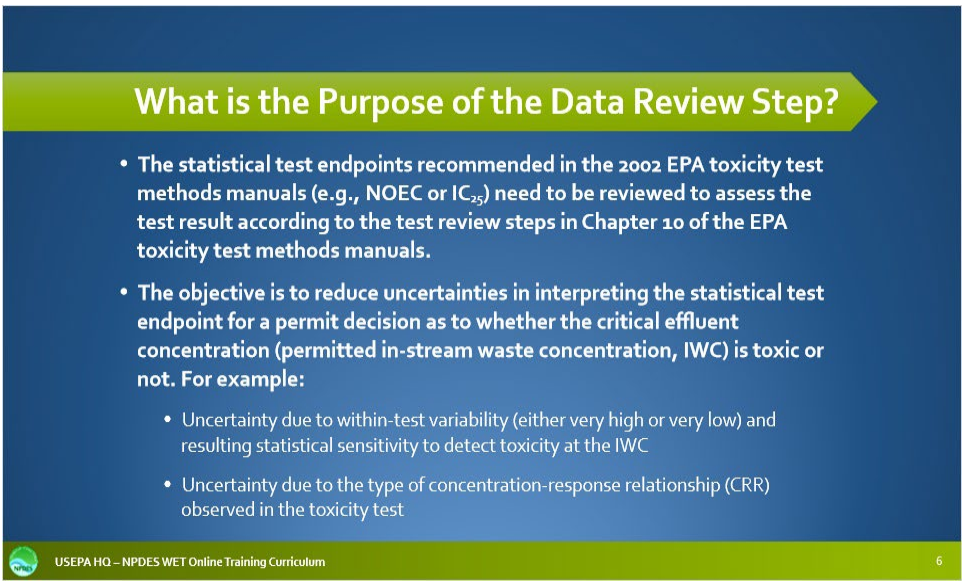
Notes:

Under the EPA 2002 toxicity test methods, toxicity data should be reviewed to determine whether test acceptability criteria, or TAC, are met, as well as sample holding times and storage requirements. In addition, the data should be reviewed to determine whether each toxicity test meets EPA's required and recommended toxicity test conditions including acceptable water quality parameter ranges, and whether the associated reference toxicant tests were completed. For this module, the assumption is made that all the required and recommended reviews were completed and that all the requirements under the EPA 2002 toxicity test methods were met. This module includes a discussion of the review of within-test variability, expressed as the PMSD, when using the recommended hypothesis statistical analysis approach for sub-lethal test endpoints. This module discusses the required review of the concentration-response relationship observed in multi-concentration toxicity testing using either the point estimate or the hypothesis statistical approach.



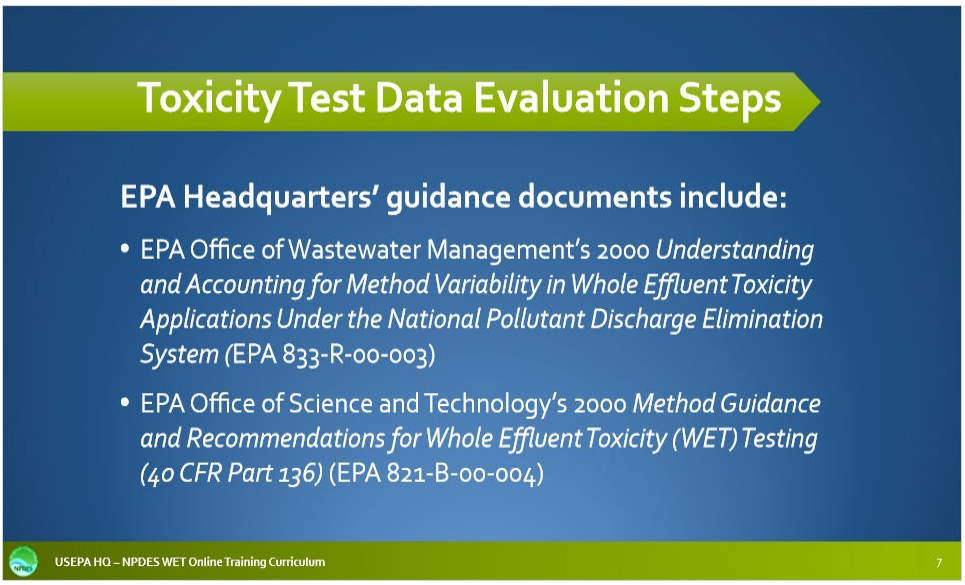
Notes:

The first step during the process of conducting toxicity testing is to collect an effluent sample according to the sample collection procedures provided in the EPA toxicity test methods. Step two is to run the toxicity tests according to the prescribed EPA toxicity test methods. Third, the organism biological responses, including mortality, and the short-term chronic sub-lethal test endpoints according to each EPA toxicity test method are recorded. Fourth, valid toxicity test data are analyzed using scientifically defensible statistical approaches, such as provided in EPA's toxicity test method appendices, that are used for the fifth and final step: to determine whether the permitted effluent's toxicity exceeds NPDES permit toxicity triggers or a WET limit. This module will discuss the interpretation of toxicity test data in step five based on review steps discussed in the EPA toxicity test methods.



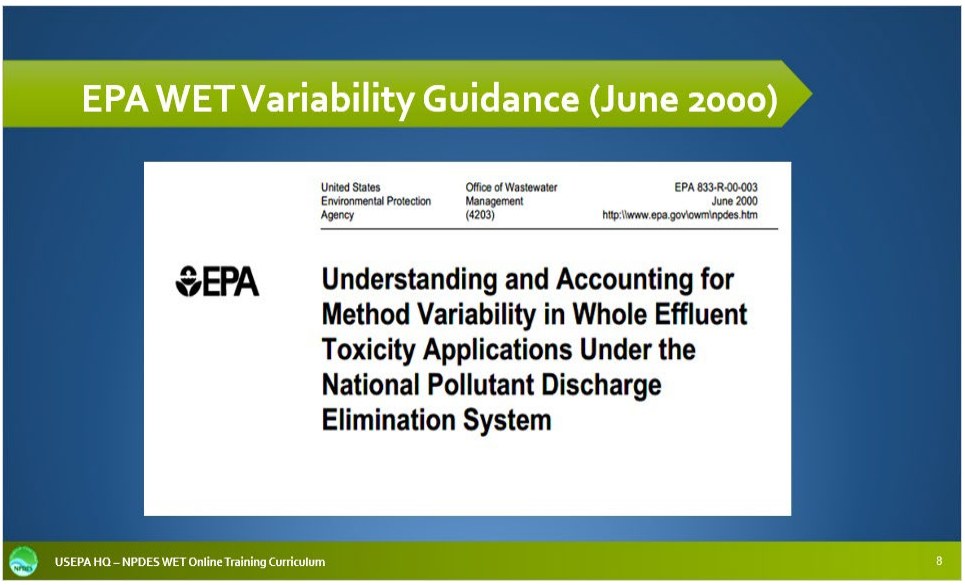
Notes:

In Chapter 10 of EPA's 2002 short-term chronic toxicity test methods, the statistical approach endpoints calculated for sub-lethal test endpoints in a valid toxicity test must be reviewed to assess the test results. When using a multi-concentration toxicity test and the hypothesis or point estimation statistical approach, the CRR must be reviewed to ensure the calculated test results are interpreted appropriately. When NPDES permits require sub-lethal hypothesis statistical test endpoints for certain EPA toxicity test methods, the within-test variability, as measured by the PMSD, must be reviewed and the allowable variability thresholds applied. The objective of these reviews is to reduce uncertainty in interpreting the statistical test endpoint for a permit decision as to whether the effluent at the critical concentration, referred to as the permitted in-stream waste concentration or IWC, is toxic or not. Uncertainty in data interpretation of a toxicity test could be due to either unusually high or low within-test variability, or due to the type of CRR observed in the toxicity test, as discussed in this module.



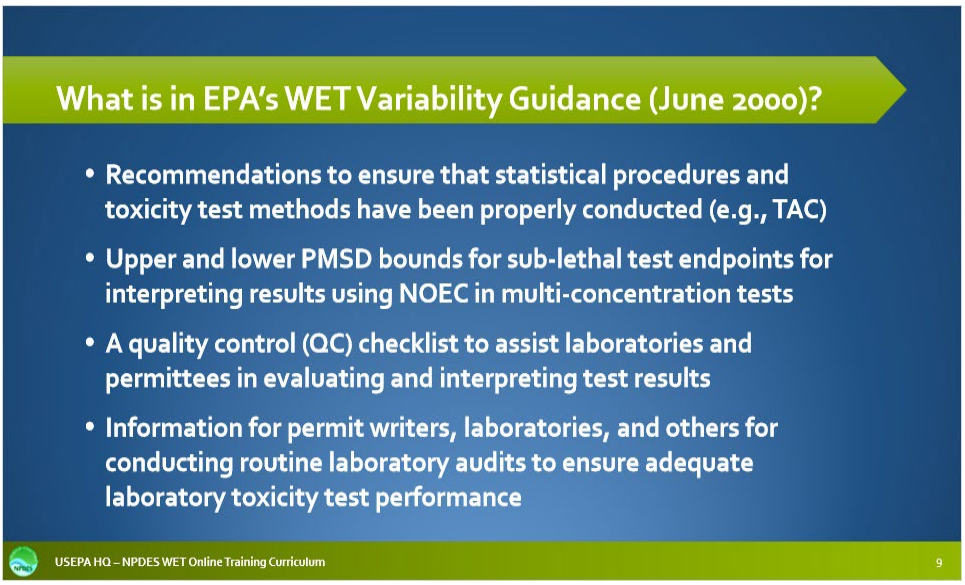
Notes:

Over the next couple of slides, we will focus on the steps used in evaluating toxicity test data. The two EPA Headquarters' guidance documents are: the Office of Wastewater Management's *Understanding and Accounting for Method Variability in Whole Effluent Toxicity Applications Under the National Pollutant Discharge Elimination System* and the Office of Science and Technology's *Method Guidance and Recommendations for Whole Effluent Toxicity (WET) Testing (40 CFR Part 136)*. Both EPA guidance documents are available in the resources tab at the top of the module and are also available on the respective EPA Headquarters offices' websites.



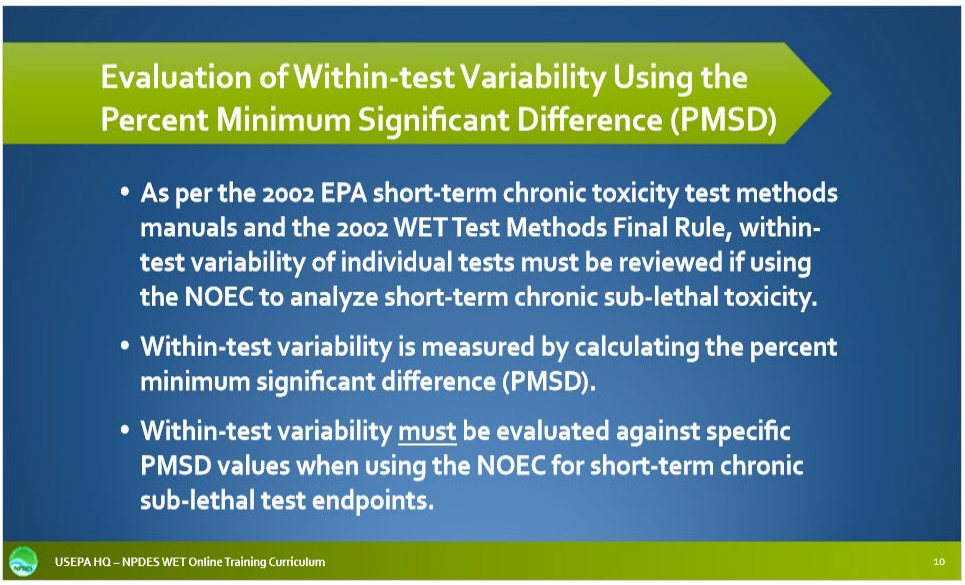
Notes:

In June of 2000, EPA's Water Permits Division in the Office of Wastewater Management released a guidance document entitled, *Understanding and Accounting for Method Variability in Whole Effluent Toxicity Applications Under the National Pollutant Discharge Elimination System*, hereafter referred to as the EPA 2000 WET Variability guidance.



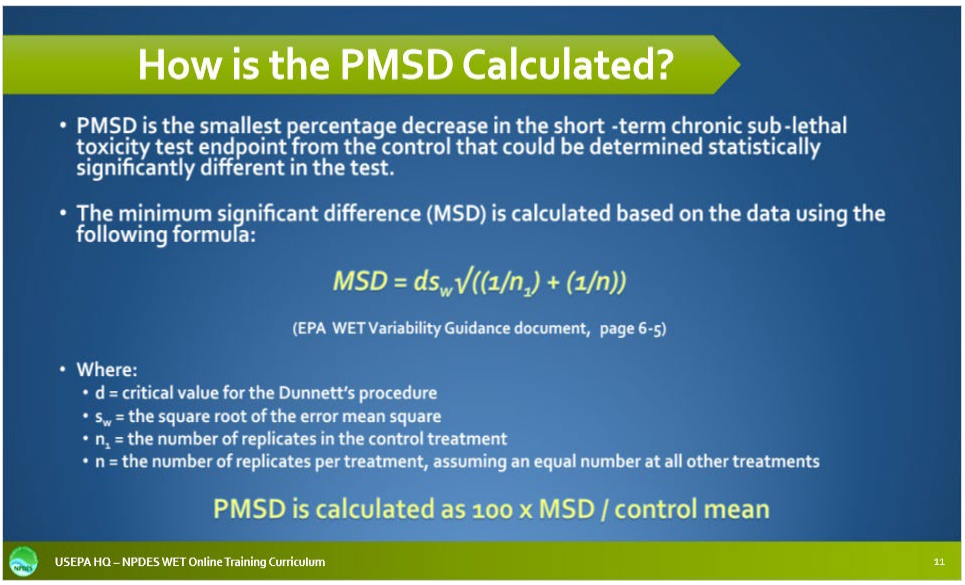
Notes:

This guidance was developed after EPA had evaluated the quality of toxicity test results generated throughout the U.S. to help permittees understand how to increase the quality of toxicity data they were generating and thereby toxicity test performance. Another important reason that EPA released this NPDES WET guidance was to provide recommendations for properly conducting the statistical analysis approaches and EPA toxicity test methods used. EPA included recommended upper and lower percent minimum significant difference, or PMSD, bounds for some EPA short-term chronic sub-lethal toxicity test method endpoints to provide guidance on acceptable within-test precision for these EPA toxicity test methods when analyzing toxicity data using a multi-concentration hypothesis statistical approach to calculate the no observed effect concentration, NOEC. These recommendations would improve the confidence with which permitting decisions regarding whether the effluent is toxic or not with respect to state aquatic life protection criteria and WET water quality standards can be made. This EPA guidance includes a quality control checklist to assist in the evaluation and interpretation of valid toxicity test results. In addition, procedures are included on how to appropriately conduct laboratory audits to help ensure that laboratory performance meets EPA toxicity test method TAC and PMSD requirements. This guidance includes a list of suggested questions that permittees should ask their laboratory to help ensure that valid, quality data are generated for their effluent samples submitted under NPDES permit applications and for WET permit compliance.



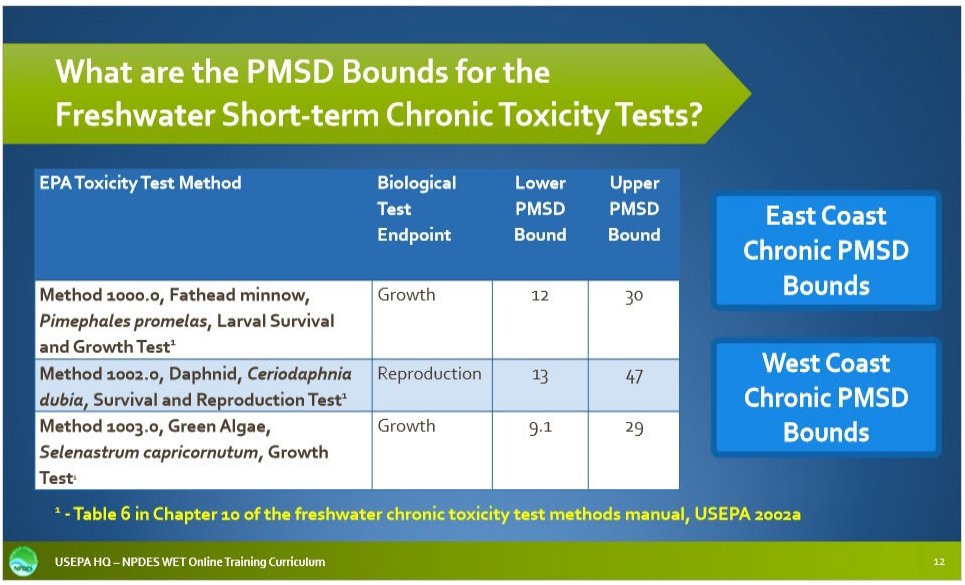
Notes:

The 2002 EPA short-term chronic toxicity test methods and the 2002 WET Test Methods Final Rule require that the within-test variability be reviewed when using the multi-concentration hypothesis statistical approach, for example the NOEC, and analyzing sub-lethal toxicity data. The derivation of the PMSD is used to measure the within-test variability. EPA has developed lower and upper PMSD bounds to which the toxicity test PMSD is compared to determine the validity of the toxicity test data. EPA did not develop upper and lower bounds for acute testing or for survival in short-term chronic tests and the PMSD is not applicable to single-concentration toxicity tests because the PMSD calculation is based on the Dunnett's t-value, which is determined using a multi-concentration toxicity test.



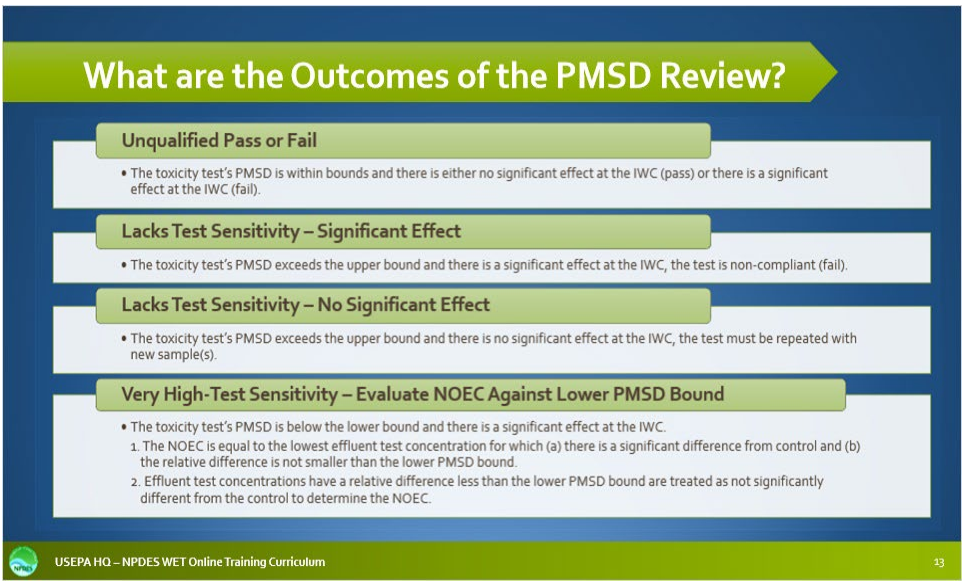
Notes:

Here we will examine how the PMSD is calculated based on the toxicity test biological response data. The PMSD is defined as the smallest percentage decrease in the sub-lethal biological test endpoint from the control that could be determined as statistically significantly different in the toxicity test. To calculate the PMSD, one must first calculate the minimum significant difference, or MSD, using the formula shown here. The MSD is calculated based on the critical value of the Dunnett's statistic as well as the square root of the error mean squares and the number of replicates in the control and effluent test concentrations. Thus, the MSD is based on the number of replicates, control performance and the power of the toxicity test. The PMSD is calculated as 100 times the MSD divided by the control mean. Commercially available toxicological evaluation software, as well as the EPA Headquarters' toxicity spreadsheet statistical tool, will generate the PMSD for toxicity test data automatically. Once the PMSD is calculated, it is compared to the established upper and lower bounds for that specific toxicity test method sub-lethal test endpoint. On the next slide we will review the established upper and lower bounds for the freshwater short-term chronic toxicity test methods and sub-lethal test endpoints.



Notes:

In the EPA 2002 short-term chronic freshwater toxicity test methods manual, Table 6 in Chapter 10 lists the lower and upper PMSD bounds for the sub-lethal hypothesis test endpoint for the freshwater short-term chronic toxicity test methods including the fathead minnow, *Pimephales promelas*, growth, *Ceriodaphnia dubia* reproduction, and the green algae, *Selenastrum capricornutum*, growth. These lower and upper PMSD bounds were determined from the 10^{th} and 90^{th} percentile, respectively, of PMSD data from the EPA's WET Inter-laboratory Variability Study from 2001. The lower PMSD bounds, which are based on the 10th percentile of national PMSD data represent the practical limit to the sensitivity of the toxicity test methods because few laboratories can achieve such precision on a regular basis, and most do not achieve it even occasionally. To assist in reviewing within-test variability, EPA recommends maintaining a control chart of PMSD values calculated for successive toxicity tests. A control chart of PMSD values characterizes the range of variability observed within a given laboratory and allows comparison of individual toxicity test PMSD values with the laboratory's typical range of variability.

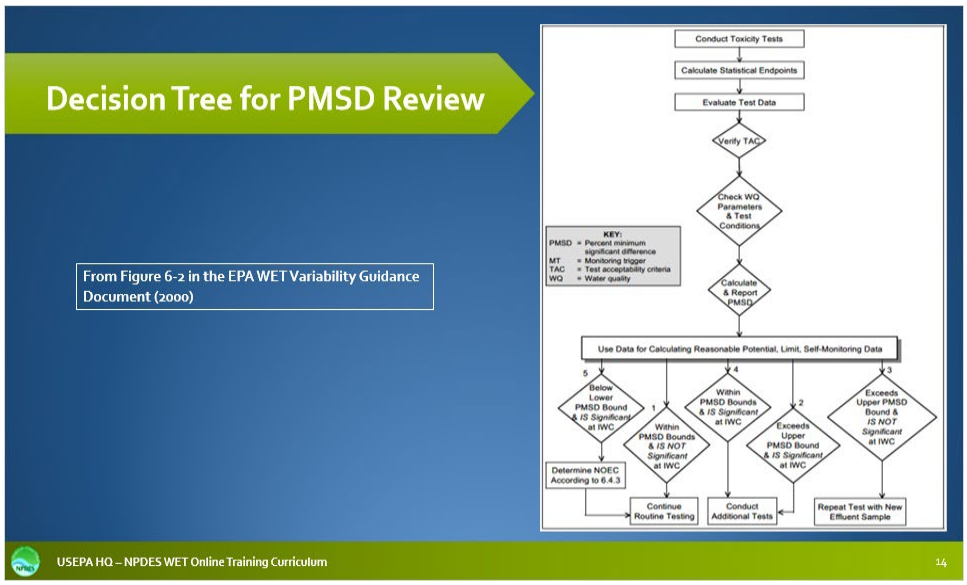


Notes:

When comparing the toxicity test PMSD with the upper and lower bounds specified by EPA and shown in the previous slide, there are four possible outcomes regarding the interpretation of the toxicity test results. If the PMSD is within the lower and upper bounds, the toxicity test result should be accepted as to whether it is a "pass" or "fail" in terms of permit compliance; that is the statistical significance of effects observed at the IWC in the toxicity test should be acceptable. Thus, if the NOEC analysis indicates that the NOEC is greater than the IWC, then the toxicity test is a "pass" and if the NOEC is less than the IWC, then the toxicity test is a "fail."
If the PMSD is greater than the upper bound for the sub-lethal biological test endpoints specified by EPA, then there are two possible outcomes: first, if the NOEC analysis indicates a significant effect at the permitted IWC, the toxicity test is considered a "fail," even though the PMSD exceeded the upper bound. This means that despite the high within-test variability, the NOEC analysis could still discern a significant effect at the IWC and therefore, the toxicity test result should be used by the permitting authority. If, however, the NOEC analysis indicates that the toxicity test is a "pass" when the PMSD exceeds the upper bound, this suggests that the within-test variability was high enough to mask a real effect of the effluent at the IWC; that is, the toxicity test lacked sensitivity due to the high within-test variability. In this case, the toxicity test result must not be accepted by the permitting authority and a new toxicity test using a new sample must be required of the permittee. If the toxicity test results persist in exceeding the upper PMSD bound, the permittee and/or the testing laboratory should be notified, and actions taken to decrease the

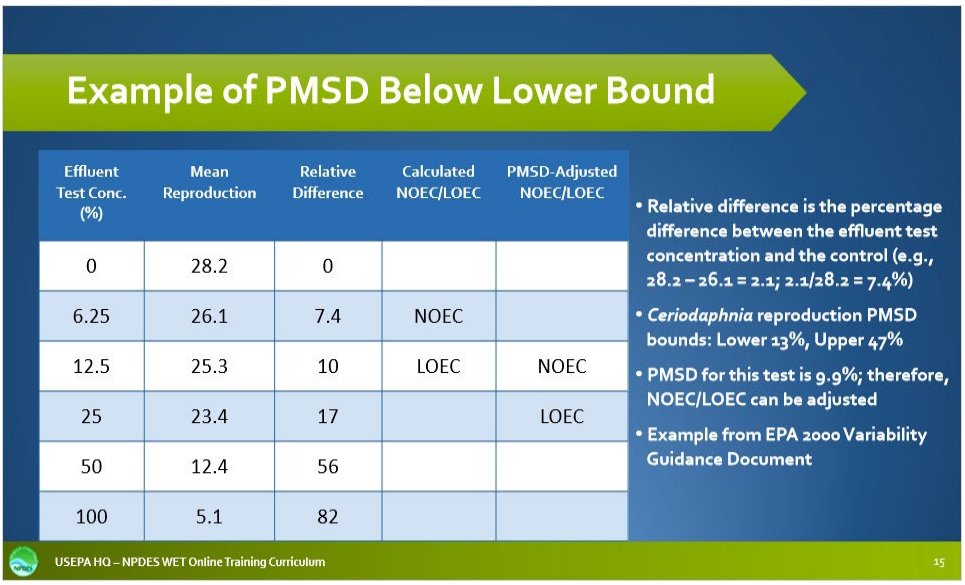
within-test variability and thereby increase the overall toxicity test performance and data quality.

The fourth possible outcome is that the test PMSD is lower than the lower bound PMSD identified by EPA. This outcome indicates that within-test variability is relatively low, which means the within-test precision is relatively high in comparison to what most laboratories have been able to achieve for the toxicity test method and biological test endpoint. Using the hypothesis NOEC statistical approach, as precision increases, a smaller difference between the control and IWC will be found to be significantly different. Therefore, the effluent has a greater probability of being noncompliant with their toxicity permit condition or that is the toxicity test will be a "fail," as the within-test variability decreases. If the PMSD for a toxicity test is less than the lower bound and the NOEC indicates no significant difference between the control and the IWC, the toxicity test should be considered a "pass" as the data are acceptable. If, however, the NOEC indicates that the toxicity test is a "fail" in this case, EPA recommends that the NOEC may be revised upward based on the lowest toxicity test concentration in which the relative difference from the control is greater than or equal to the lower PMSD bound and there is a significant effect. In this case, lower toxicity test concentrations as compared to the revised NOEC are not considered significantly different. The revised NOEC should be compared with the permit triggers or WET limit to determine compliance with the permit toxicity conditions.



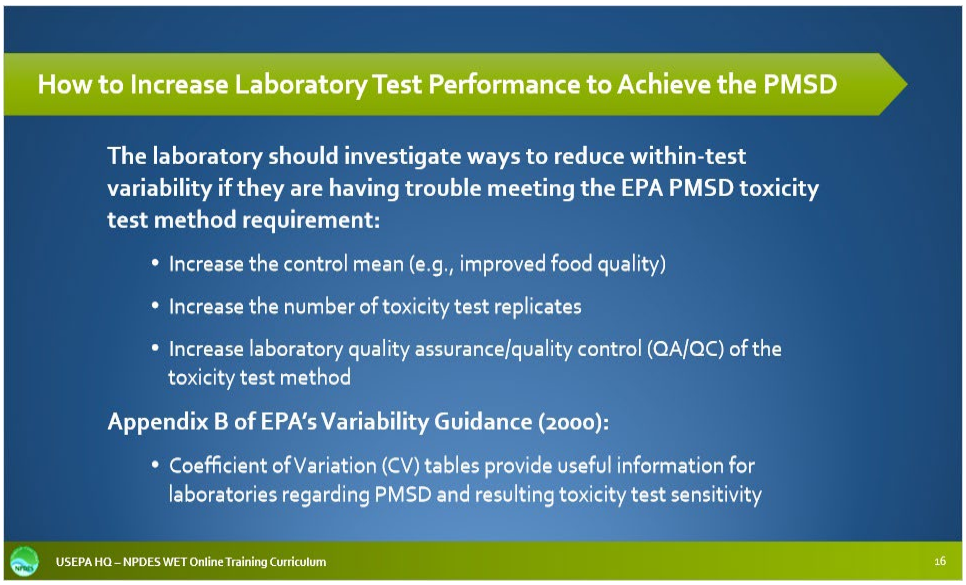
Notes:

To review, if the PMSD is within the lower and upper bounds, the NOEC result should be used, depending on results of the concentration-response relationship review step, and the toxicity test is identified as a “pass” or “fail” accordingly. If the PMSD is higher than the upper bound and the NOEC result indicates a “fail,” the toxicity test result should be considered a “fail” and accelerated toxicity testing should be conducted, if accelerated monitoring was required in the permit. If, however, the NOEC analysis indicated a “pass” when the upper PMSD bound is exceeded, the toxicity test result should not be considered conclusive and a new toxicity test using a new sample, or samples should be required. If the PMSD for the toxicity test is less than the lower PMSD bound as recommended by EPA for the toxicity test method and biological test endpoint, and the toxicity test result is a “pass,” then that result should be considered conclusive and toxicity testing continued as required in the permit. If, however, the NOEC analysis indicates that the test is a “fail” then in this case, the NOEC can be adjusted upward as noted in the previous slide and compliance evaluated based on the upward-revised NOEC following the procedure shown in the next slide and discussed in EPA’s WET Variability Guidance in section 6.4.3.



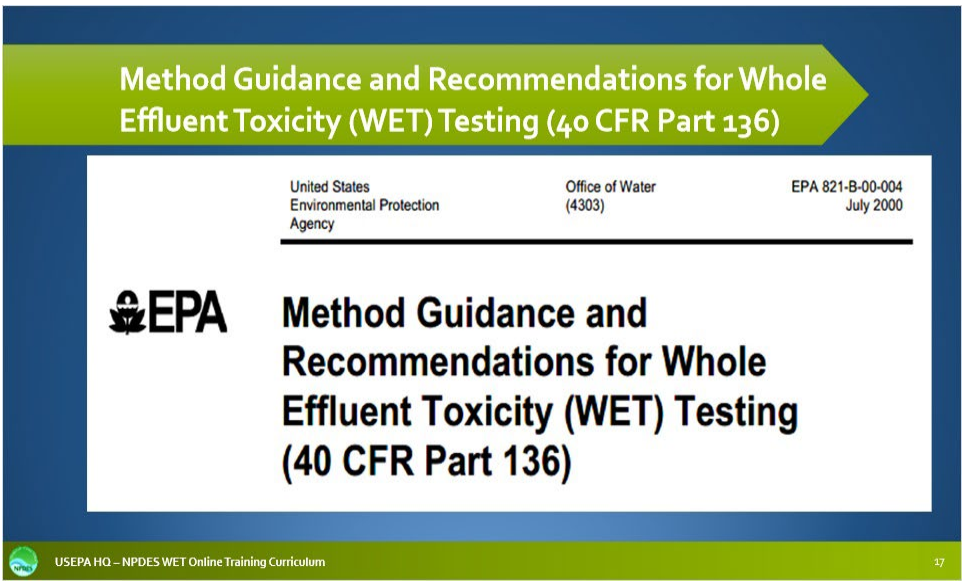
Notes:

This slide shows an example of how the NOEC can be adjusted upward when the test PMSD is lower than the lower PMSD bound as specified by the EPA 2000 WET Variability Guidance. In this example using *Ceriodaphnia dubia* reproduction data, the toxicity test PMSD of 9.9 percent is lower than EPA's lower PMSD bound of 13 percent for this toxicity test endpoint. The NOEC should be adjusted as explained previously because of the low PMSD. In this example, the lowest test concentration at which the relative difference is at or above the lower bound PMSD is 25 percent effluent (17 percent difference), which now becomes the lowest observed effect test concentration or LOEC. Therefore, the revised NOEC is now 12.5 percent effluent.



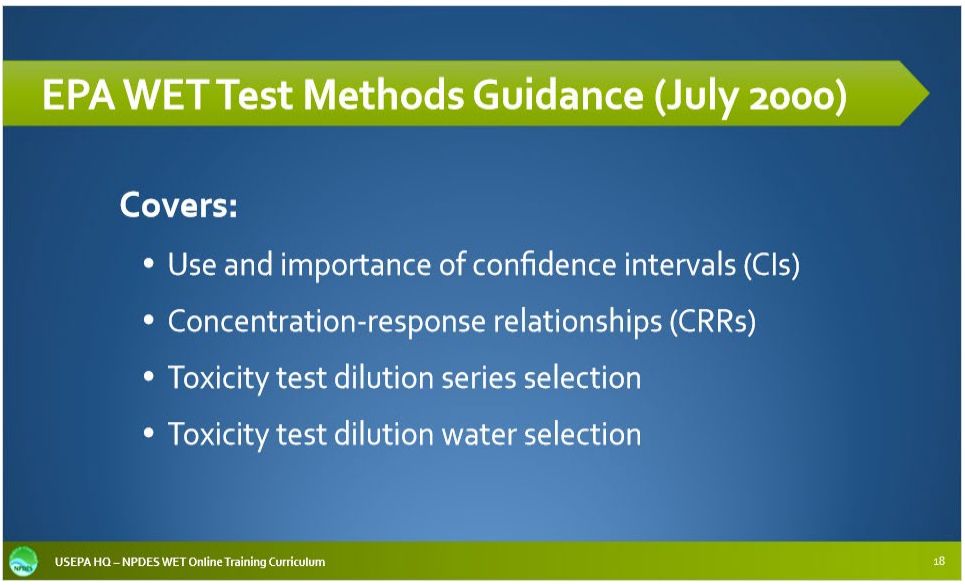
Notes:

If a laboratory is having trouble meeting the EPA PMSD toxicity test method requirement, or is frequently experiencing high control variability within a toxicity test, or high variability in a given biological test endpoint between reference toxicant tests, EPA's 2000 WET Variability Guidance discusses ways that laboratories can: reduce their within-test variability due to laboratory performance, develop and implement a rigorous quality assurance/quality control, QA/QC program, increase toxicity test organism performance, use toxicity test organism food of the appropriate quality, and, if need be, increase the number of toxicity test replicates for each toxicity test concentration and control treatments within a toxicity test. Remember that the number of replicates per toxicity test concentration given in the EPA toxicity test methods is a required minimum number. This means that a laboratory could increase the number of replicates to reduce within-test variability, and thereby increase laboratory performance and resulting data quality. Other recommendations provided in EPA's 2000 WET Variability Guidance include appendices that discuss appropriate reference toxicants and reference toxicant testing procedures, as well as a system that laboratories can use to track biological test endpoint-specific coefficient of variation, or CV. The CV should be reported as part of the control chart developed for each species tested. Appendix B also offers guidance on the range of CVs that should be observed for each EPA toxicity test species and biological test endpoint.



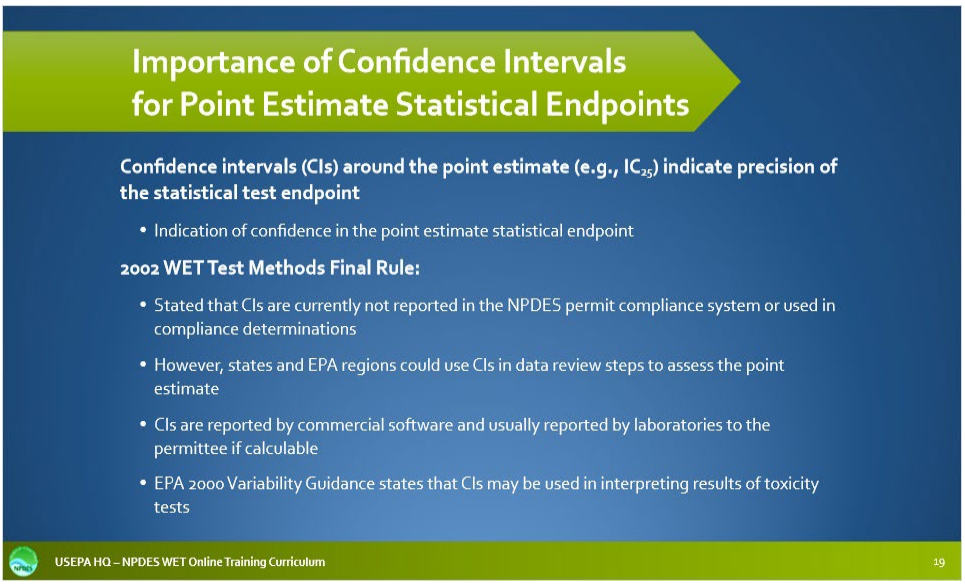
Notes:

EPA's Office of Science and Technology's Engineering and Analysis Division's 2000 Method Guidance and Recommendations for Whole Effluent Toxicity (WET) Testing (40 CFR Part 136), hereafter referred to as EPA 2000 WET Method Guidance, is another EPA guidance document that provides useful information for permittees and laboratories regarding toxicity test data interpretation.



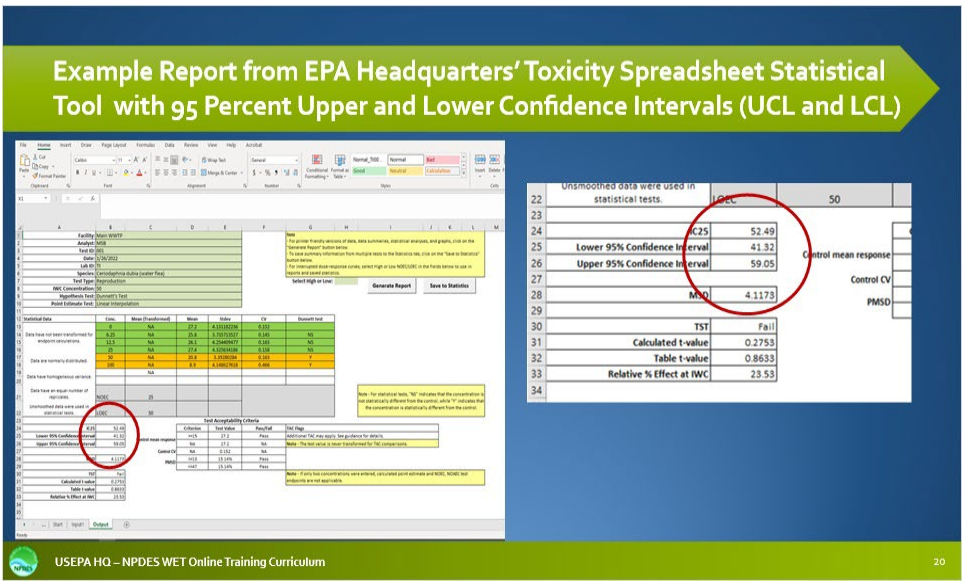
Notes:

This guidance discusses the importance of confidence intervals, or CIs, when interpreting point estimate statistical approach endpoints and how to properly apply CIs in point estimate analyses. Another topic of interest in this guidance includes examples of different types of concentration-response relationships, or CRRs, and how to evaluate data from those CRRs. In addition, this guidance discusses recommended effluent dilution series for different effluent scenarios and how to select the proper test dilution water for NPDES WET permit monitoring. As explained in the NPDES Testing Methods for Whole Effluent Toxicity module and in the NPDES WET Permit Conditions, Permit Language and Technical Considerations module, these factors can have a profound effect on the statistical approach endpoints reported and the confidence in those statistical approach endpoints.



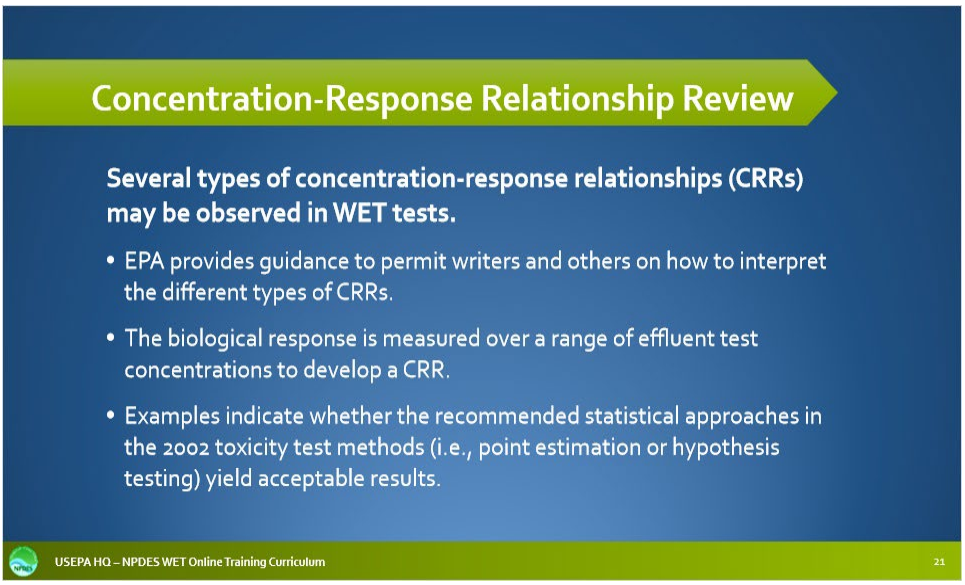
Notes:

The statistically calculated CIs determined in the point estimate analysis are an indication of the precision of the statistical approach endpoint and the confidence in the mean point estimate identified. CIs are statistically derived from the mean and the range of biological responses observed in each replicate of each test concentration in a toxicity test. When the range of CIs between the lower 95 percent confidence level and the upper 95 percent confidence level is large, then there is less confidence in the mean point estimate identified, including the IC_{25}. Although CIs are calculated by most toxicity statistical analysis software, including EPA Headquarters' toxicity spreadsheet statistical tool, if CIs can be calculated, the 2002 WET Test Methods Final Rule states that CIs are currently not required to be reported in the NPDES compliance system or used in toxicity compliance determinations. However, states and EPA regions may use CIs in the review of toxicity data to assess the point estimate result as per the 2000 EPA WET Methods Guidance. On the next slide we will look at a statistical analysis reporting sheet that includes the CIs surrounding the statistical point estimate endpoint results.



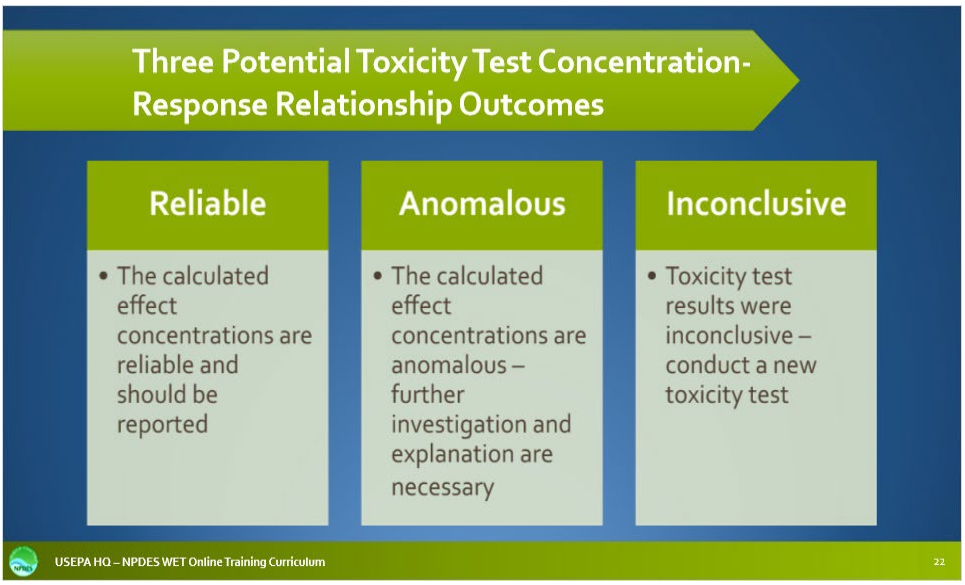
Notes:

Here we see an example statistical report from EPA Headquarters' toxicity spreadsheet statistical tool. Under the summary of the point estimate statistical analysis, there are multiple percent effect values reported from an IC_5 or five percent effect to an IC_{50} or 50 percent effect. As noted in the NPDES WET Statistical Analysis and Toxicity Data Review module, the effect level used for NPDES permit compliance in short-term chronic *Ceriodaphnia* toxicity testing is 25 percent or an IC_{25}. The IC_{25} is reported as 52.49 percent effluent with a lower 95 percent confidence level of 41.32 percent effluent and an upper 95 percent confidence level of 59.05 percent effluent. Thus, based on the within-test variability observed in this toxicity test, statistically there is a 95 percent certainty that the actual mean IC_{25} is between 41.32 and 59.05 percent effluent. Under the current NPDES regulations, CIs are not used in determination of toxicity permit compliance; but if the in-stream waste concentration of this effluent is within the range of the 95 percent CIs, there is the potential that the IC_{25} is less than the IWC and that the effluent is not compliant with the toxicity permit condition, even though the toxicity test would be a "pass" based on the mean IC_{25} of 52.49 percent effluent identified in this example.



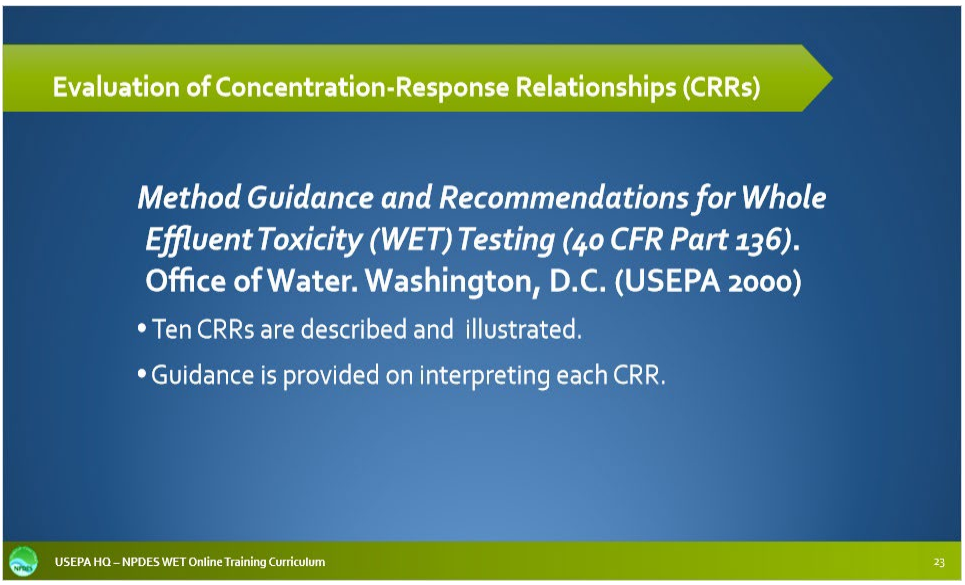
Notes:

Another toxicity test data review step that is required for multi-concentration tests under the EPA 2002 toxicity test methods manuals and outlined in the EPA 2000 WET Method Guidance is review of the concentration-response relationship or CRR. EPA's 2002 toxicity test methods require toxicity tests to be conducted using a minimum of five effluent test concentrations and a control concentration. This toxicity test design is necessary to calculate the NOEC hypothesis statistical approach endpoint and a point estimate statistical approach endpoint, such as IC_{25}. EPA recognized that in multi-concentration toxicity tests, many types of CRRs may be observed depending on the effluent test concentrations or dilution series used in a toxicity test, as well as organism health, and many other factors. The 2000 EPA WET Method Guidance discusses the review of the CRR in a multi-concentration toxicity test to assist in interpreting test results using either the NOEC or IC_{25} to assess these statistical approach test endpoints.



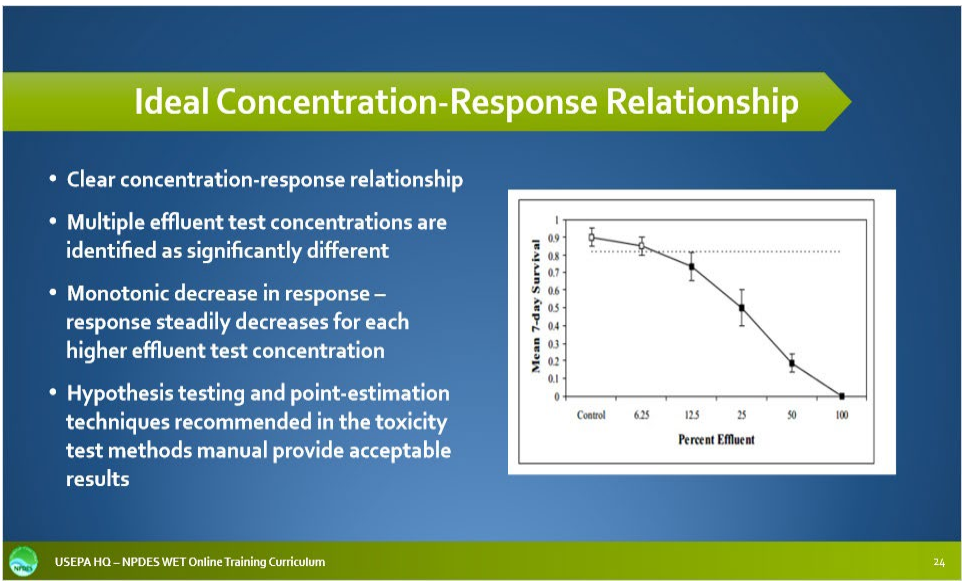
Notes:

As we noted in the previous slide, EPA's 2000 WET Method Guidance describes different potential concentration-response relationships and how they should be evaluated to determine if results should be used in NPDES permitting decisions. Three main outcomes of a concentration-response relationship review are identified: 1) the calculated effect concentration should be used, 2) the calculated effect concentration is anomalous and further investigation and explanation are necessary before it should be used, and 3) the toxicity test results are inconclusive, and a new toxicity test should be initiated with a newly collected sample. These three outcomes of a concentration-response relationship review will be examined in more detail in this module and how they are determined based on a CRR review.



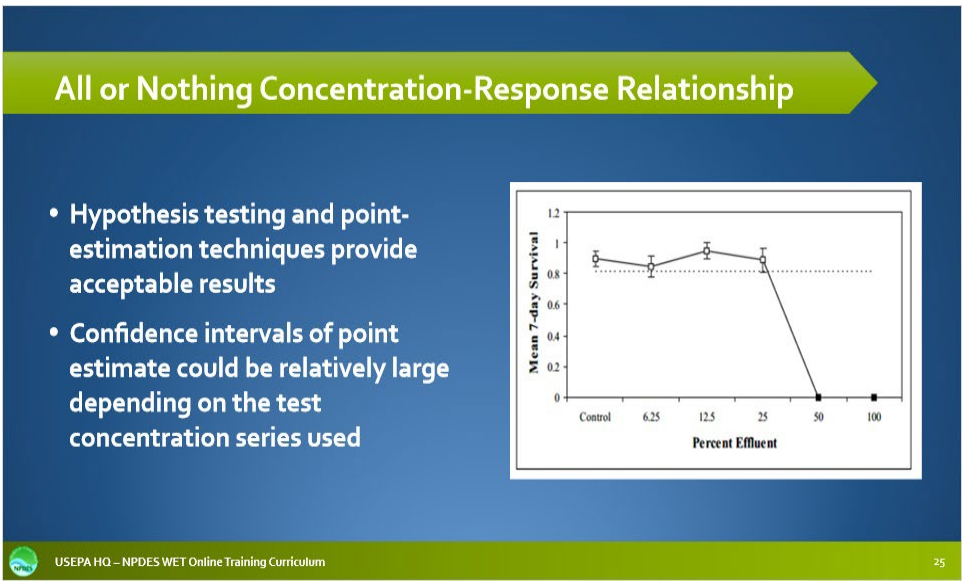
Notes:

In EPA's 2000 WET Method Guidance, ten of the most common CRRs observed in toxicity testing are described and illustrated. The guidance also details how to interpret each CRR and ways to assess each statistically derived test endpoint. Over the next ten slides we will briefly examine each of these example CRRs and how EPA's 2000 WET Method Guidance provides recommendations for assessing each statistical test endpoint.



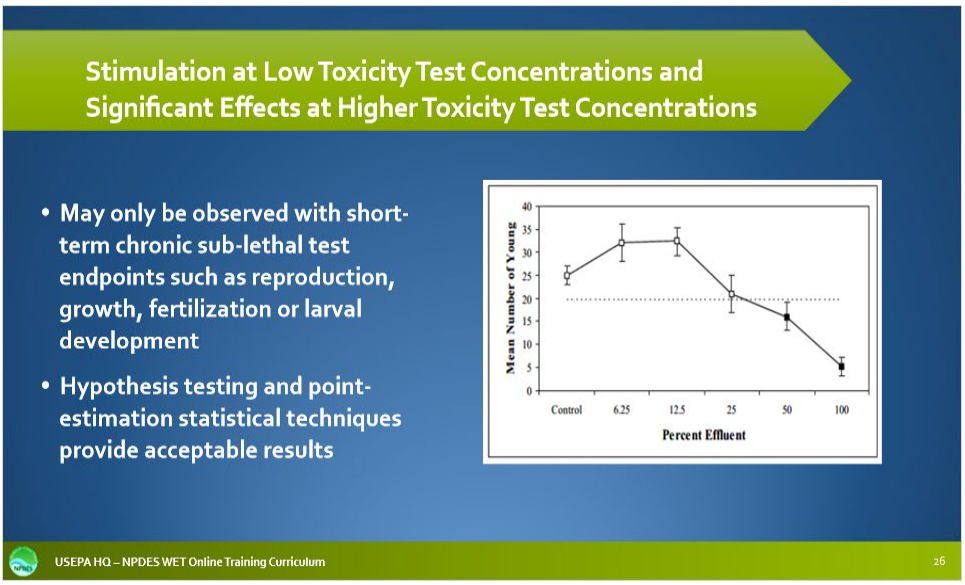
Notes:

The first concentration response relationship examined in the EPA 2000 WET Method Guidance document is what is termed an "ideal concentration-response relationship." The concentration response is clear; multiple test concentrations are identified as being significantly different from the control and the concentration response is monotonically decreasing; that is, the organism response, in this case survival, decreases as the percent effluent increases. This CRR is commonly produced when testing chemicals where the concentration series can be well-designed with appropriately chosen test concentrations that bracket the chemical's range of toxicity. As the 2002 WET Test Methods Final Rule emphasizes, such an ideal CRR is not necessarily expected nor is it required for toxicity tests. When this type of CRR is encountered the statistical results from both hypothesis testing and point estimation statistical approaches recommended in the 2002 toxicity test methods manuals provide acceptable results.



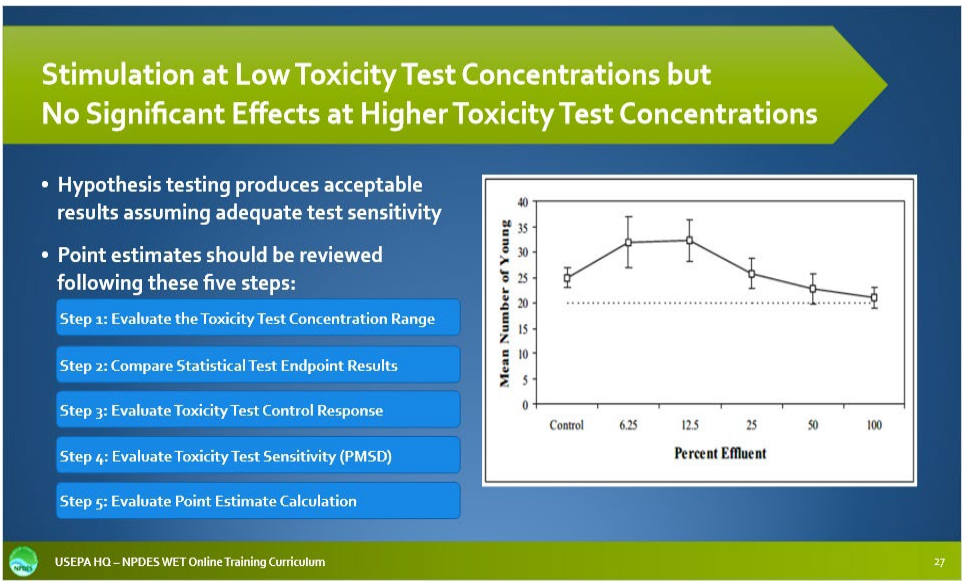
Notes:

The CRR illustrated here represents an all or nothing response and is very common in toxicity test data. This CRR is characterized by a transition from no significant effect at one effluent test concentration to a complete effect, for example 100 percent mortality, at the next higher test concentration. This CRR also produces acceptable results using both hypothesis and point estimation statistical approaches recommended in the EPA 2002 toxicity test methods manual. This CRR is indicative of a steep CRR and under these circumstances, the precision of the estimate may be improved by closer spacing of toxicity test concentrations or the addition of intermediate toxicity test concentrations in future testing.



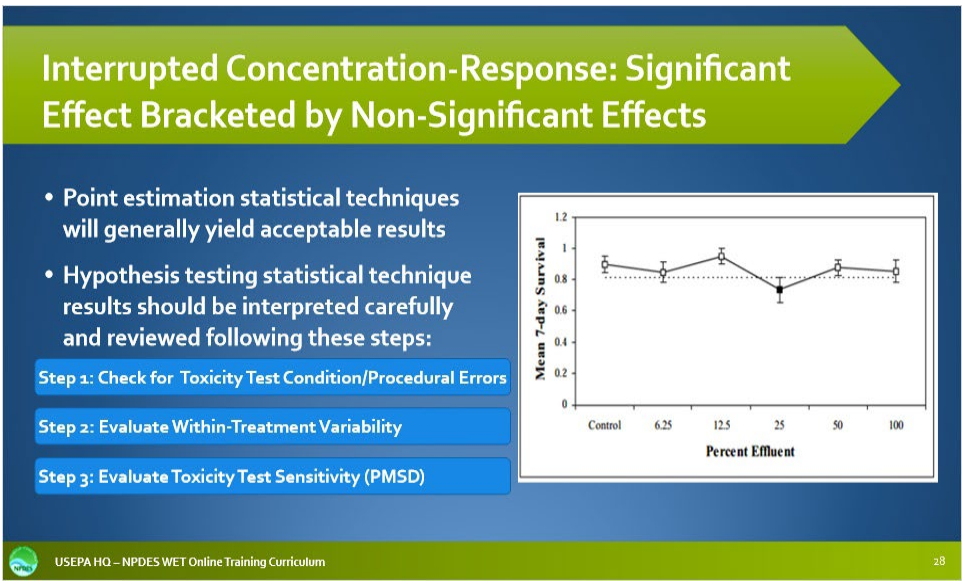
Notes:

This CRR is characterized by a stimulation at low test concentration but significant effects at higher test concentrations. This CRR is considered nonmonotonic and is typically found with sub-lethal biological test endpoints such as reproduction, growth, fertilization, or larval development. For instance, *Ceriodaphnia* reproduction may increase relative to the controls at low test concentrations of an effluent due to an increased nutrient and food availability and decrease relative to the control at higher test concentrations where concentrations of potential toxicants outweigh the benefits of increased nutrients. This CRR pattern, while nonmonotonic, still produces acceptable results using both hypothesis testing and point estimation statistical approaches recommended in the EPA 2002 toxicity test methods manual.



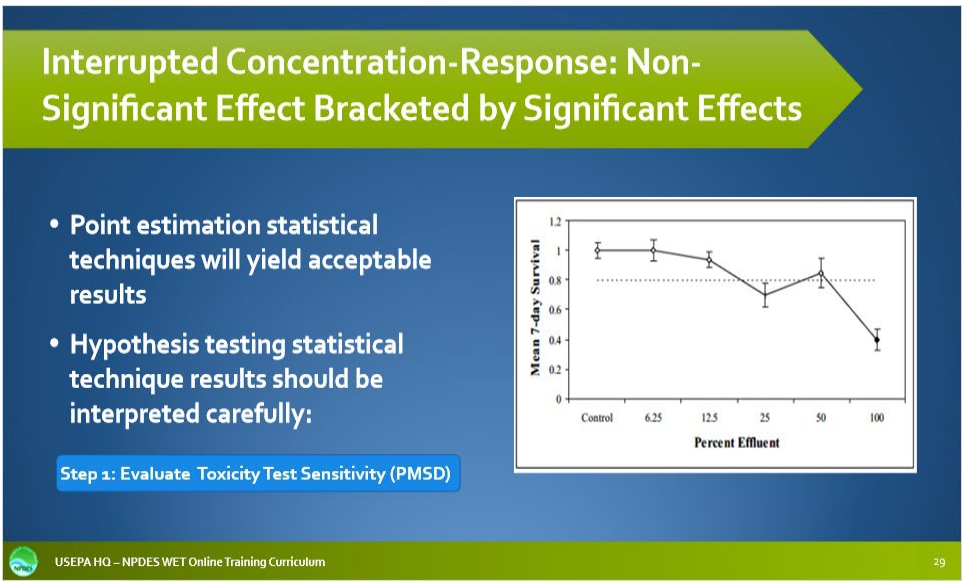
Notes:

This CRR is similar to the previous example in that stimulation is observed at lower test concentrations, but in this case, higher test concentrations do not produce significant effects. In this situation, hypothesis testing statistical approaches should produce acceptable results, assuming that adequate test sensitivity is achieved. Results from point estimation statistical approaches should be interpreted carefully when this CRR is encountered, because the inhibition concentration procedure may produce effect concentrations, particularly IC_{25}s, that indicate toxicity at toxicity test concentrations where the response is comparable to the control response. In cases where the responses at the low toxicity test concentrations are much higher than in the control, the smoothing process, as detailed in the 2002 EPA toxicity test methods, may result in a large upward adjustment in the control mean. This can lead to an IC_{25} result that is less than the highest test concentration, even though the highest test concentration was not statistically different from the control treatment, even if a percent difference of less than 25 percent was observed between the control response and the response at the highest test concentration. If this type of CRR is encountered, the following review steps should be taken in addition to standard toxicity test review procedures: 1) evaluate the toxicity test concentration range, 2) compare hypothesis testing and point estimates results, 3) evaluate the control response, 4) evaluate the toxicity test sensitivity (PMSD), and 5) evaluate the point estimate calculation. Click on the blue buttons on this slide for a detailed explanation of each review step.



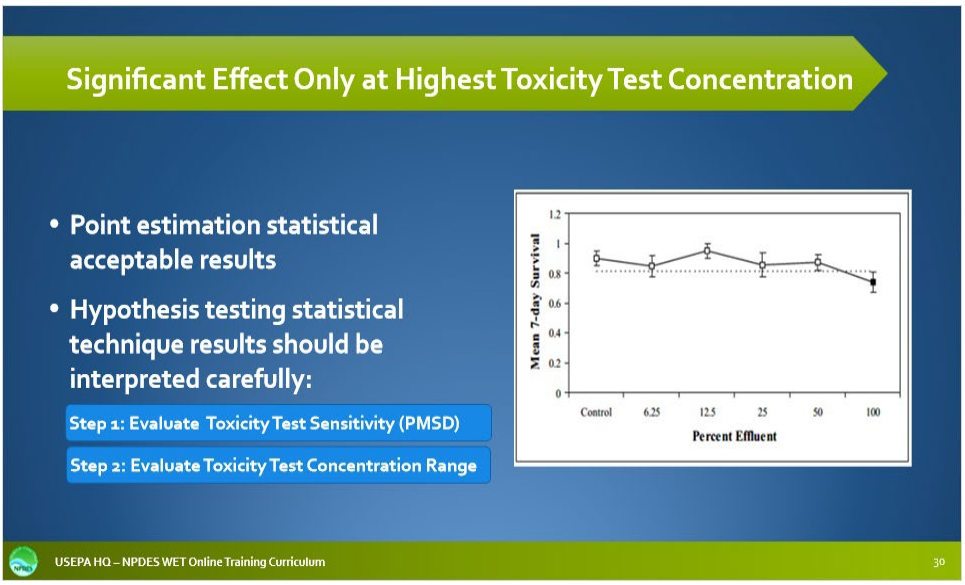
Notes:

Next, we will examine two CRRs that are considered interrupted concentration-responses. In this first CRR, the response relationship is characterized by a single test concentration showing a significant difference from the control while the adjacent higher and lower test concentrations do not differ significantly from the control. When this response relationship is encountered, point estimation statistical approaches will generally yield acceptable results, but the hypothesis testing statistical approach results should be interpreted carefully. The identification of the NOEC and LOEC using the hypothesis testing statistical approaches was intended for situations in which as the toxicity test concentration gets higher, the response, that is the effect, also increases. In this situation where the CRR is non-monotonic, the identification of the NOEC and LOEC values may be compromised. Review of the hypothesis testing statistical approach results should include the following several steps: 1) checking the toxicity test conditions or for procedural errors, 2) evaluating within-treatment variability, and 3) evaluating the toxicity test sensitivity (PMSD). Click on the blue buttons on this slide for a detailed explanation of each review step.



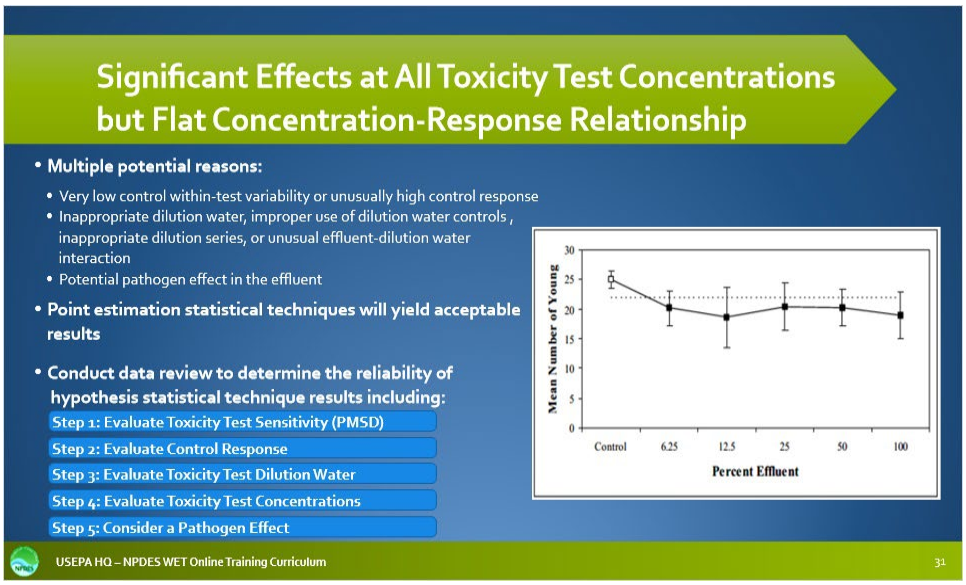
Notes:

This response relationship depicted here is similar to the previous response relationship in that the CRR is non-monotonic or interrupted, however, this CRR is characterized by two or more test concentrations showing a significant difference from the control while an intermediate test concentration does not differ significantly from the control. When this CRR is encountered, the point estimation statistical approach will yield acceptable results, but the hypothesis testing statistical approach results should be interpreted carefully. As mentioned for the previous CRR, the identification of the NOEC and LOEC values is affected when the CRR is non-monotonic. Click on the blue button on this slide for a detailed explanation of the review step.



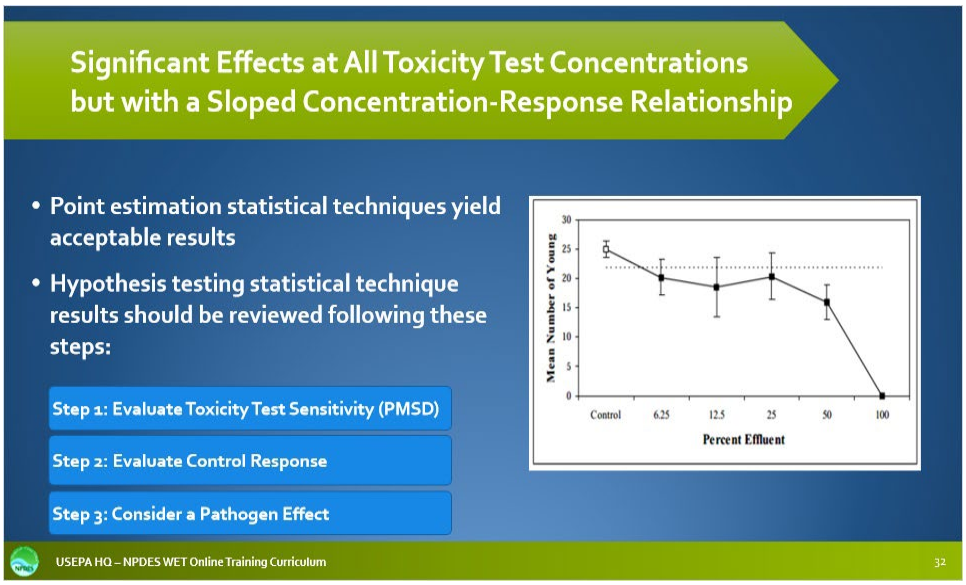
Notes:

This response relationship is characterized by only the highest test concentration producing a significantly different response from the control. The results determined by the point estimation statistical approach should be acceptable. The hypothesis testing statistical approach results may also be determined to be acceptable following an evaluation of toxicity test sensitivity. If the response relationship depicted here is observed, where a significant effect occurs only at the highest test concentration, the calculated toxicity test sensitivity (PMSD) and the test concentration range should be evaluated. Click on the blue buttons on this slide for a detailed explanation of each review step.



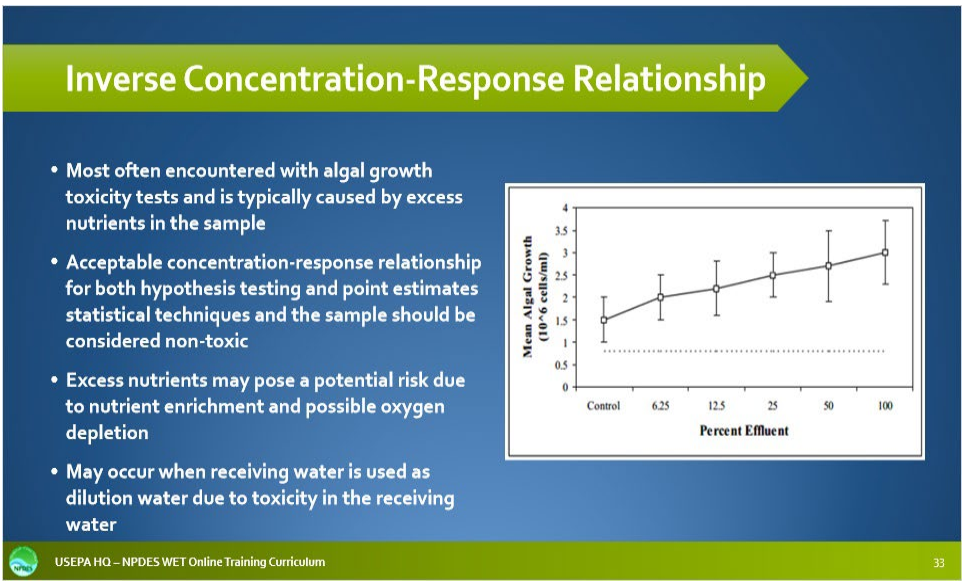
Notes:

In this CRR, all the test concentrations produce a response that is significantly different from the control response, but a clear concentration-response relationship cannot be determined. The results determined by point estimation statistical approaches should be assumed to be acceptable. Hypothesis testing statistical approach results may be found to be acceptable following an evaluation of the toxicity test sensitivity, the control response, test dilution water or the test concentration series, and a potential pathogen effect. This CRR could be due to: 1) an extremely low variability in the control, 2) an unusually high control response, 3) an inappropriate dilution water and/or improper use of dilution water controls, 4) an inappropriate test dilution series, 5) potential pathogen effects in the effluent, or 6) an unusual effluent-dilution water interaction. The review actions described in the blue buttons slides should be taken to determine a cause for this concentration-response relationship and to subsequently determine the validity of the calculated results.



Notes:

This CRR is similar to the CRR identified in the previous slide except that a decreasing CRR can be identified at higher toxicity test concentrations. Point estimation statistical approaches will yield acceptable results, but results determined with hypothesis testing statistical approaches should be interpreted carefully and the cause for significantly different effects at low test concentrations should be investigated as described for the previous CRR with significant effects at all test concentrations but no clear concentration-response relationship.

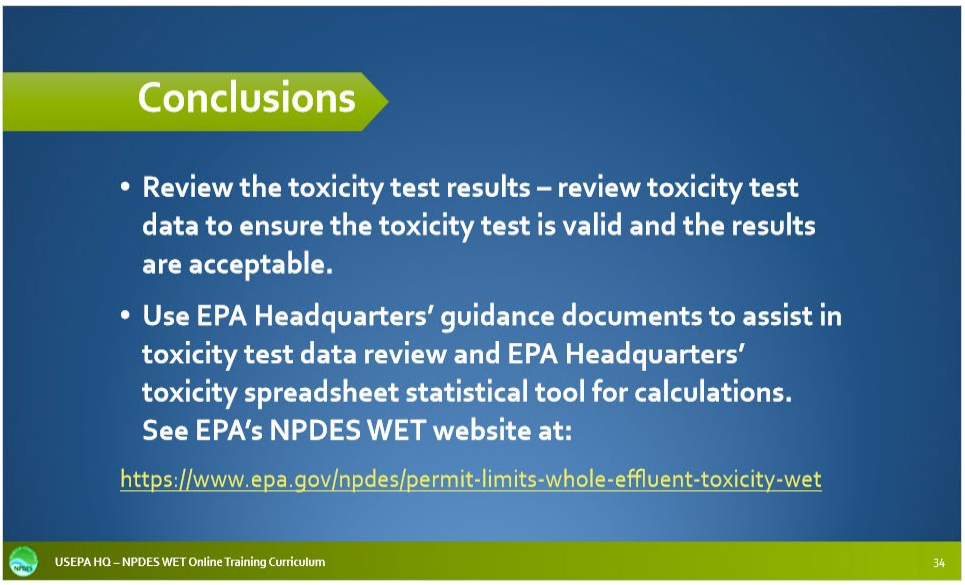


Notes:

This CRR is characterized by a relationship in which adverse effects decrease with increasing toxicity test concentration. This situation is most often encountered in algal growth tests and is typically caused by excess nutrients in the sample. When this type of CRR is encountered the statistical results from both hypothesis testing and point estimation statistical approaches recommended in the EPA 2002 toxicity test methods manuals provide acceptable results. The sample should be considered non-toxic since the direction of the CRR indicates decreasing adverse effects. It should be noted that while the sample is considered non-toxic, the presence of excess nutrients still may pose a potential risk to the environment due to nutrient enrichment and oxygen depletion.

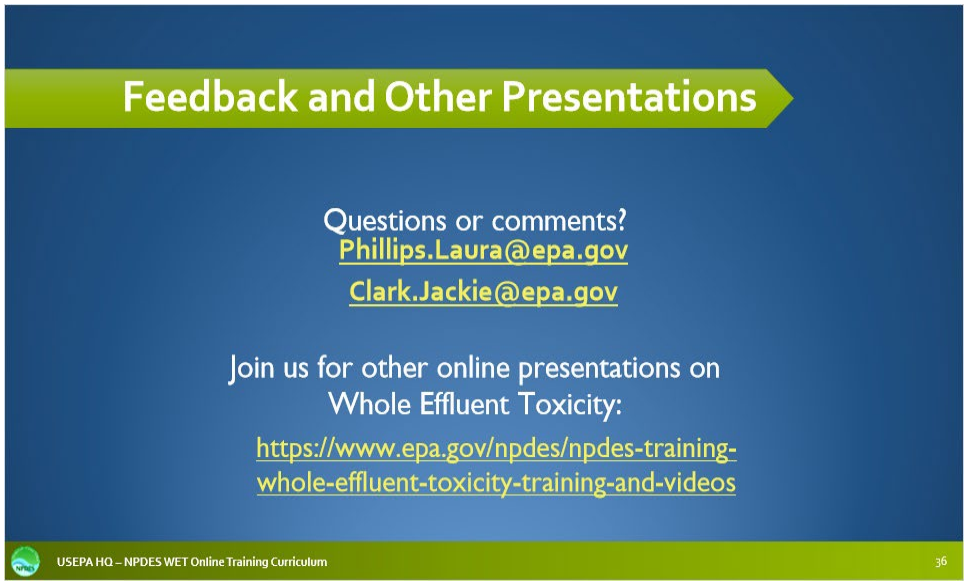
An inverse CRR also may occur in toxicity tests other than algal growth assays when the dilution water used is a receiving water or synthetic water adjusted to approximate the receiving water characteristics. In such situations, the inverse CRR can result from toxicity in the receiving water or the limitation of necessary components, for example hardness, in the receiving water or adjusted synthetic water. Under such circumstances, the objective of the toxicity test should be evaluated, see Chapter six of EPA's 2000 Method Guidance. If the objective of the toxicity test is to determine the toxicity of the sample in the natural receiving water, then these results indicate no toxicity in the sample. If the objective of the toxicity test is to determine the absolute presence of toxicity in the sample, the sample should be tested using a newly collected sample using a standard synthetic dilution water. Toxicity or limiting components in the receiving water or adjusted synthetic

water may mask the presence of low-level toxicity in the sample, making the absolute determination of toxicity in the sample difficult.



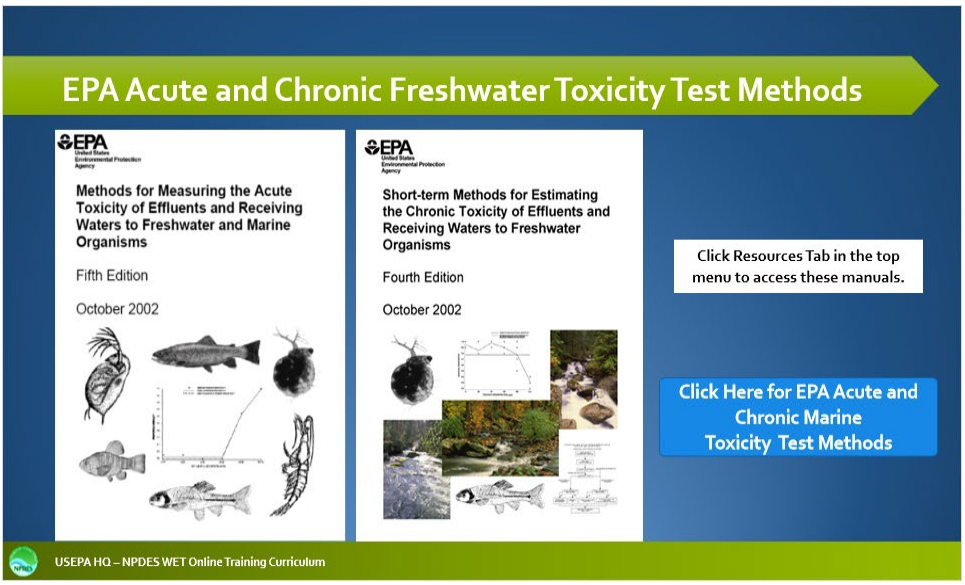
Notes:

In conclusion, the review steps outlined in this module have been developed by the EPA to ensure that toxicity test data is valid and toxicity test results are acceptable for both permitting development decisions and for the determination of compliance with the NPDES permit. The NPDES permit should incorporate these review steps as part of the process for reviewing toxicity test data and the EPA guidance documents are available for assistance. EPA Headquarters has also developed a toxicity spreadsheet statistical tool for calculation of toxicity statistical approach endpoints including the NOEC and IC_{25}. EPA Headquarters' toxicity spreadsheet statistical tool also calculates the PMSD for comparison to the developed lower and upper bounds, as well as the confidence intervals surrounding the point estimate. EPA has developed guidance and the EPA Headquarters' toxicity spreadsheet statistical tool to assist in the review of toxicity test data but when needed assistance should be requested through the state, EPA region, or EPA Headquarters.



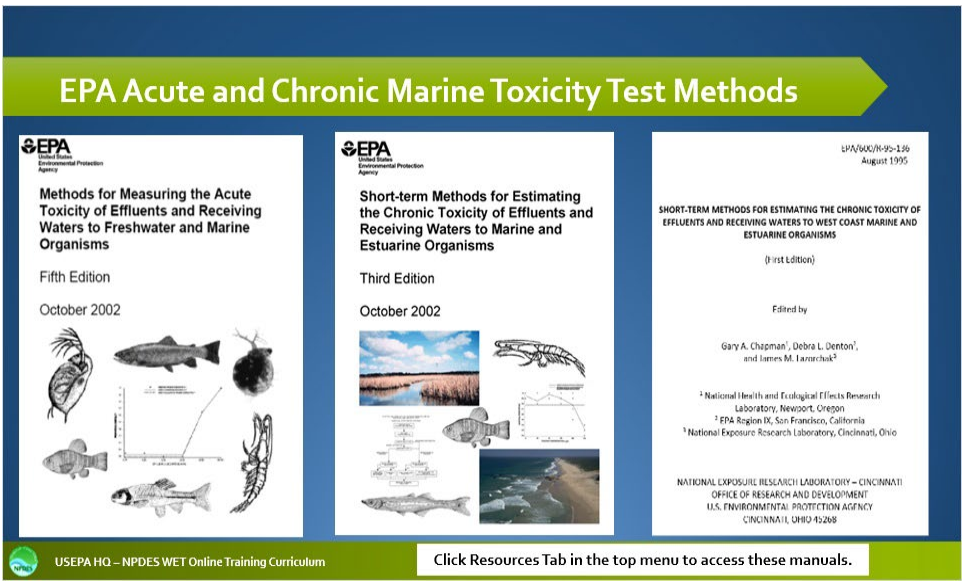
Notes:

Thank you for joining us for this EPA's NPDES Whole Effluent Toxicity training presentation. We hope that you have enjoyed it!
If you have any questions or comments on this or any part of the EPA's NPDES WET online training curriculum, click on the email address given on this slide to send a message to Laura Phillips or Jackie Clark, EPA Headquarters NPDES WET Coordinators.
Remember, you will find all the EPA's NPDES WET online training presentations, under the EPA's NPDES training section found on the Office of Wastewater Management's NPDES website.
See you next time!



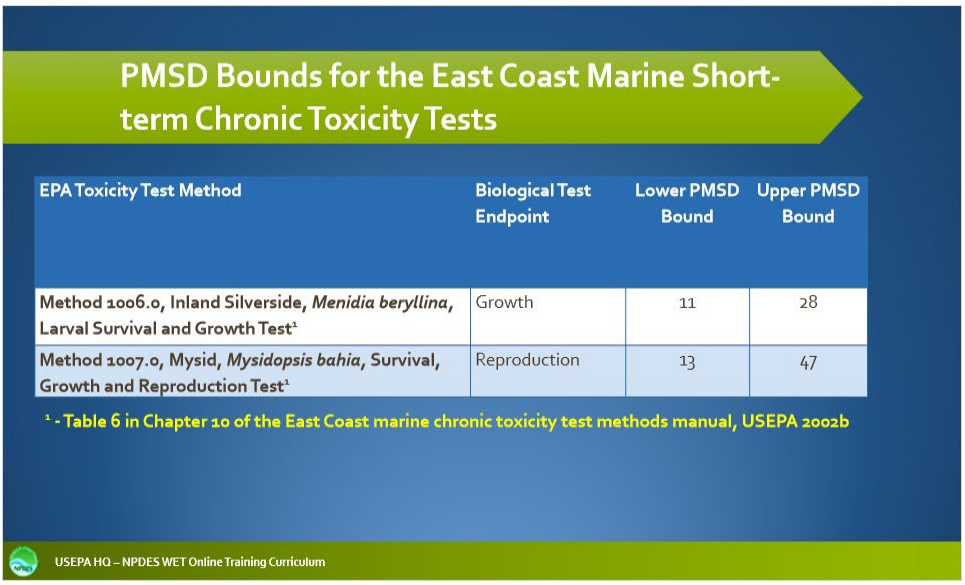
Notes:

The module presented here examines EPA's freshwater acute toxicity test methods entitled, *Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms*, Fifth Edition, EPA-821-R-02-012, hereafter acute toxicity test methods. In addition, this module provides EPA's short-term chronic freshwater toxicity test methods entitled *Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms*, Fourth Edition, EPA-821-R-02-013, hereafter chronic toxicity test methods.



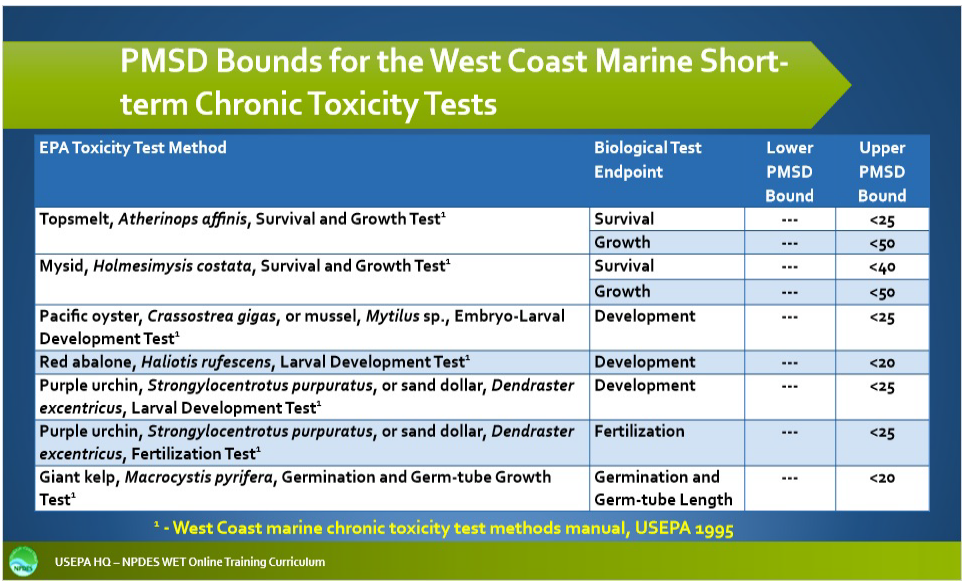
Notes:

This course also provides an opportunity to view EPA's acute marine toxicity test methods entitled *Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms*, Fifth Edition, EPA-821-R-02-012, hereafter acute toxicity test methods; short-term chronic marine toxicity test methods used by states on the Atlantic Ocean or Gulf of Mexico entitled *Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms*, Third Edition, EPA-821-R-02-014, hereafter East Coast chronic toxicity test methods; or short-term chronic marine toxicity test methods used by states on the Pacific Ocean entitled *Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to West Coast Marine and Estuarine Organisms*, First Edition, EPA-600-R-95-136, hereafter West Coast chronic toxicity test methods.



Notes:

The PMSD bounds for the EPA East Coast marine toxicity test methods have only been calculated for the inland silverside fish and the mysid shrimp toxicity tests. The lower and upper PMSD bounds for inland silverside, *Menidia beryllina*, fish growth chronic sub-lethal biological test endpoint are 11 and 28 percent, respectively. The chronic sub-lethal biological test endpoint for the mysid shrimp, *Americamysis bahia* formerly *Mysidopsis bahia*, toxicity test method has lower and upper PMSD bounds of 11 and 37 percent, respectively.



PMSD Bounds for the West Coast Marine Short-term Chronic Toxicity Tests

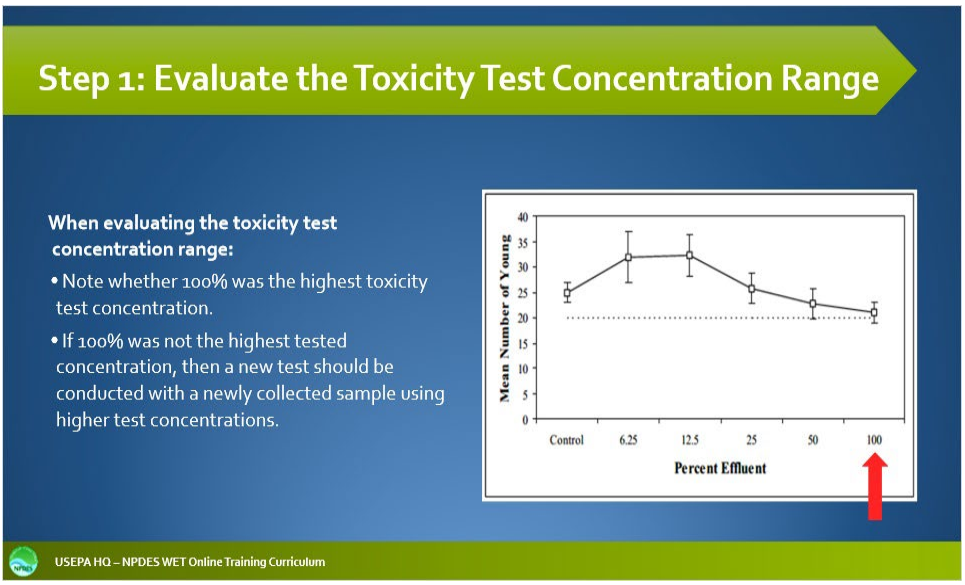
EPA Toxicity Test Method	Biological Test Endpoint	Lower PMSD Bound	Upper PMSD Bound
Topsmelt, *Atherinops affinis*, Survival and Growth Test[1]	Survival	---	<25
	Growth	---	<50
Mysid, *Holmesimysis costata*, Survival and Growth Test[1]	Survival	---	<40
	Growth	---	<50
Pacific oyster, *Crassostrea gigas*, or mussel, *Mytilus* sp., Embryo-Larval Development Test[1]	Development	---	<25
Red abalone, *Haliotis rufescens*, Larval Development Test[1]	Development	---	<20
Purple urchin, *Strongylocentrotus purpuratus*, or sand dollar, *Dendraster excentricus*, Larval Development Test[1]	Development	---	<25
Purple urchin, *Strongylocentrotus purpuratus*, or sand dollar, *Dendraster excentricus*, Fertilization Test[1]	Fertilization	---	<25
Giant kelp, *Macrocystis pyrifera*, Germination and Germ-tube Growth Test[1]	Germination and Germ-tube Length	---	<20

[1] - West Coast marine chronic toxicity test methods manual, USEPA 1995

USEPA HQ – NPDES WET Online Training Curriculum

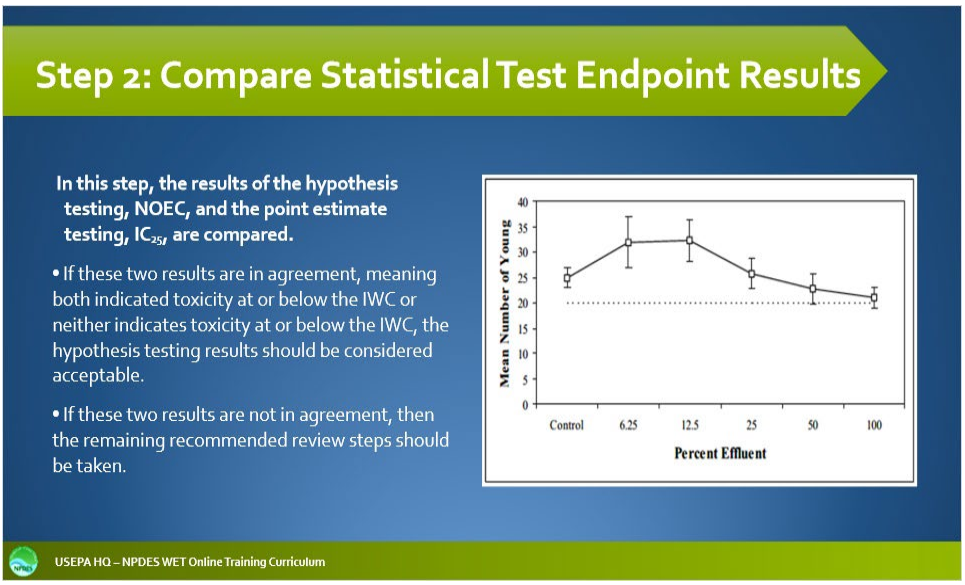
Notes:

The PMSD bounds for EPA West Coast chronic marine toxicity test methods have only been calculated for the upper bound. The upper bound for the PMSD must be less than the specific PMSD species value. For the topsmelt, *Atherinops affinis*, survival and growth test, the PMSD for survival is 25 percent and the PMSD for growth is 50 percent. The mysid shrimp, *Holmesimysis costata*, survival and growth PMSDs are 40 and 50 percent, respectively. The Pacific oyster, *Crassostrea gigas*, and mussel, *Mytilus* sp., embryo-larval development, as well as the purple sea urchin, *Strongylocentrotus purpuratus*, and sand dollar, *Dendraster excentricus*, embryo development and fertilization PMSD chronic sub-lethal biological test endpoints are all 25 percent. The red abalone, *Haliotis rufescens*, larval development and the giant kelp, *Macrocystis pyrifera*, germination and germ-tube length PMSD chronic sub-lethal biological test endpoints are all 20 percent.



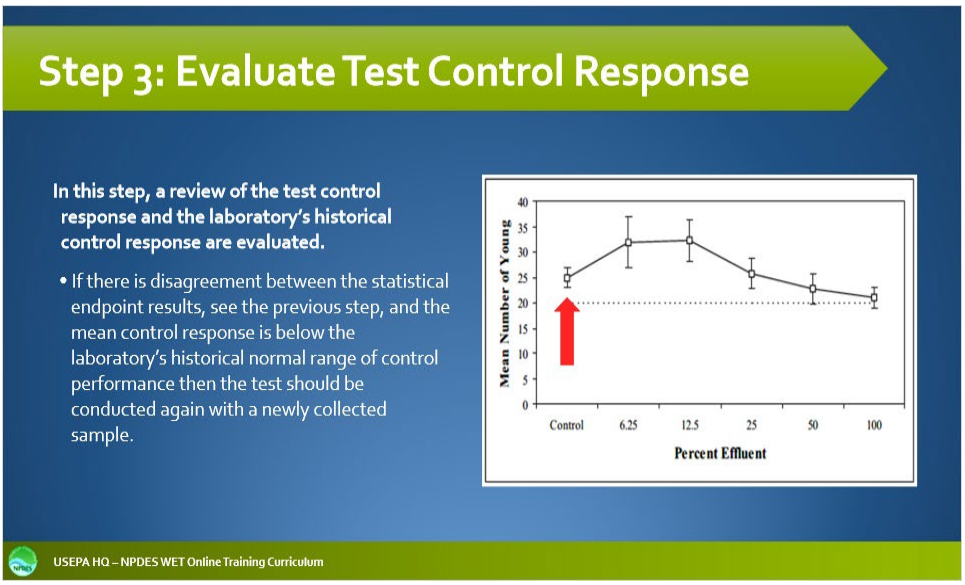
Notes:

In step one, when evaluating the concentration response, it should be noted whether the highest test concentration in the test was less than 100 percent (or the highest achievable effluent test concentration for marine tests), and if it was less than 100 percent then the toxicity test should be repeated with a newly collected sample using higher test concentrations to establish an acceptable CRR. However, this may not be necessary if the permit WET limit is set at much lower than 100 percent and test results indicate no toxicity at the permit limit level and at least one test concentration above the WET permit limit.



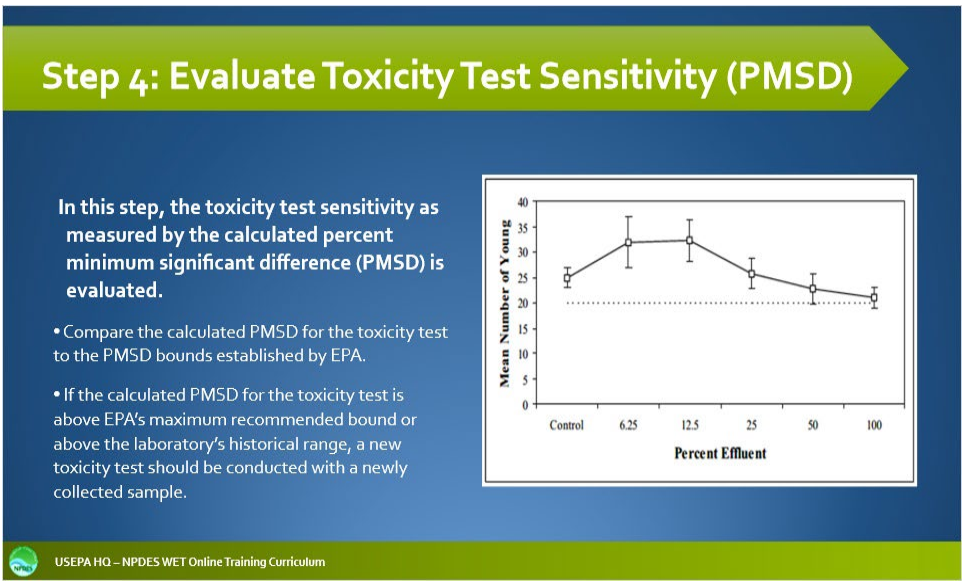
Notes:

In this step when comparing the results of the hypothesis testing and the point estimates, if there is agreement between the NOEC and the IC_{25} for tests producing this CRR, such that both values indicate toxicity at or below the IWC or neither value indicates toxicity at or below the IWC, the results should be considered acceptable. If, however, the NOEC indicates no toxicity at the IWC but the IC_{25} is calculated as less than the IWC, the remaining recommended review steps should be taken.



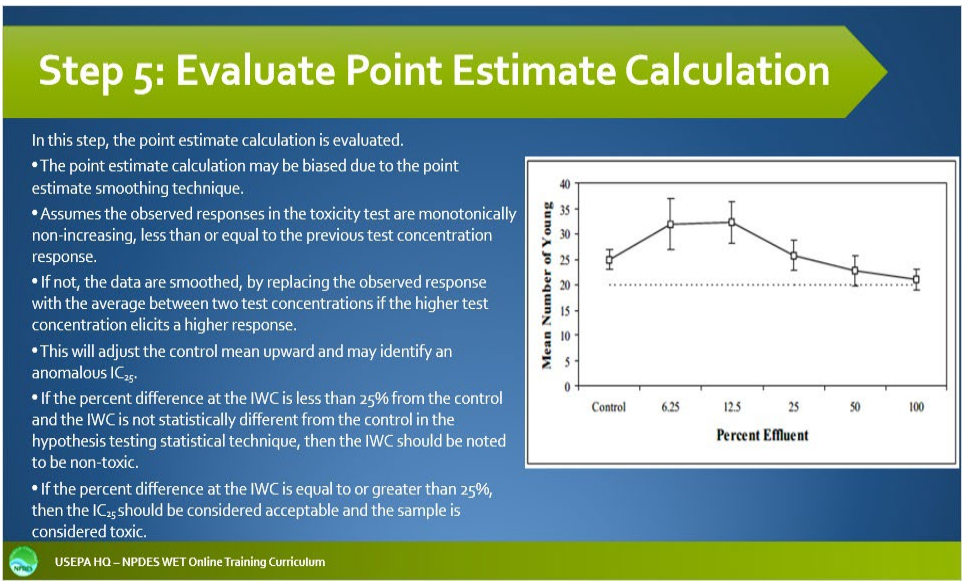
Notes:

Step three is a review of the control response in the toxicity test as well as compared to the laboratory's historical control response. If this CRR is encountered, and there is disagreement between the NOEC and IC_{25}, see step 2, and if the mean control response is well below the laboratory's normal range of control performance; the toxicity test should be conducted again with a newly collected sample even if the minimum TAC were met in the toxicity test.



Notes:

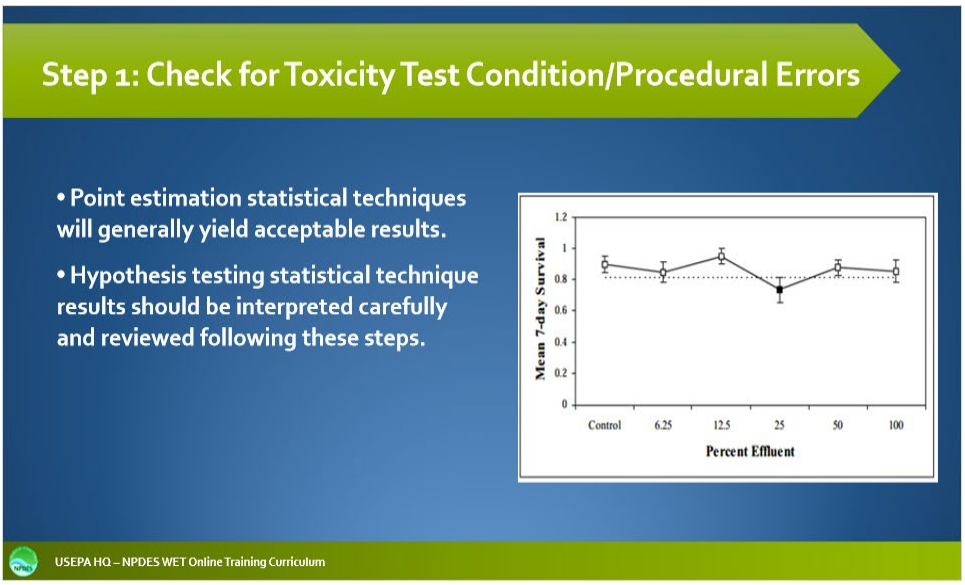
Step four is an evaluation of the test sensitivity using the percent minimum significant difference that we reviewed earlier in this module. Disagreement between NOEC and IC_{25} could be due to low toxicity test sensitivity. By comparing the percent minimum significant difference (PMSD) in the toxicity test to the PMSD bounds previously discussed and to the laboratory's historical toxicity test sensitivity performance, it may be noted that the PMSD is above the maximum recommended bound or above the laboratory's typical range, and the toxicity test should be conducted again with a newly collected sample.



Notes:

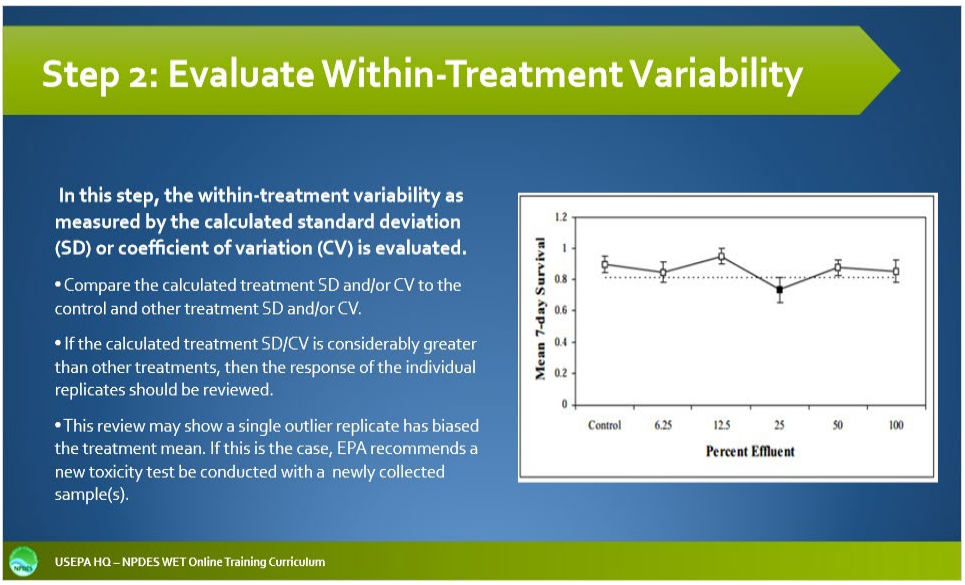
In this review step an evaluation of the point estimate calculation is conducted. If this CRR is encountered and it has been determined that it is not due to poor control performance or low-test sensitivity, then the discrepancies between the NOEC and the IC_{25} may be due to bias from the point estimate smoothing statistical approaches. The point estimate calculation assumes that the observed responses in the toxicity test are monotonically non-increasing, where the mean response for each higher test concentration is less than or equal to the mean response for the previous test concentration, and that the data follow a linear response function. When using linear interpolation and a non-monotonic CRR is observed, to account for the assumption that the data are monotonically non-increasing and follows a linear response function, the data are smoothed. Smoothing is used when non-monotonic concentration responses are observed and replaces the observed response with the average between two test concentrations if the higher test concentration elicits a higher response. By smoothing the data, this ensures a monotonically non-increasing and linear response. In cases like this CRR, where the responses at lower test concentrations are higher than in the controls, the smoothing process may result in a large upward adjustment in the control mean. To determine if this is the case, calculate the percent difference in response between the control and IWC. If this percent difference is less than 25 percent, then the response at the IWC is not statistically significantly different from the control response and the calculated IC_{25} of less than the IWC should be noted as anomalous and the effluent determined to be non-toxic at the IWC. If the observed

percent difference is equal to or greater than 25 percent, then the calculated IC_{25} should be considered acceptable.



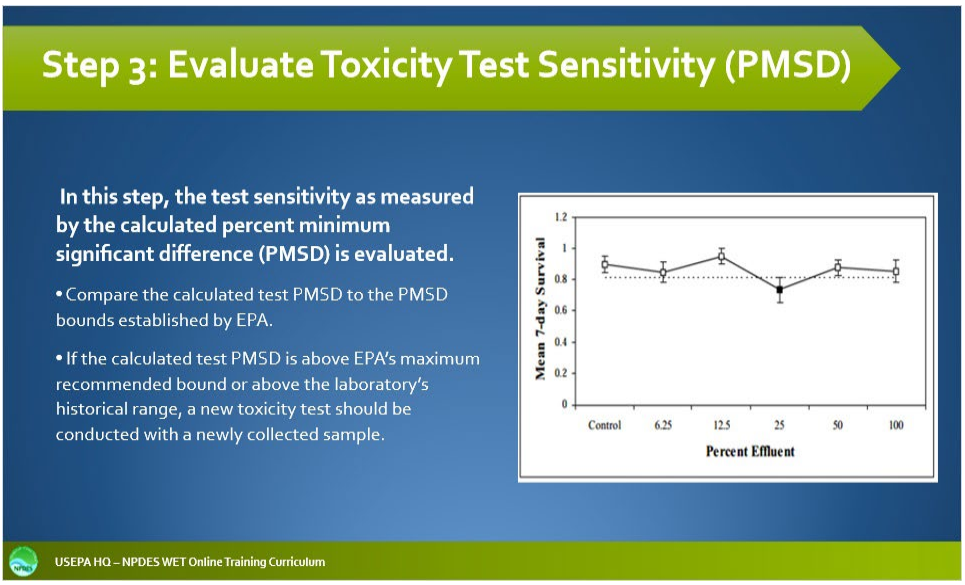
Notes:

The first review step under this CRR is to review for test conditions or procedural errors, such as pH, dissolved oxygen, conductivity or temperature excursions occurring in isolated toxicity test replicates. Other procedural errors could be due to the failure to properly randomize test organisms or test chambers. If the toxicity test conditions or procedural errors are identified, it is recommended that the toxicity test be conducted again with a newly collected sample.



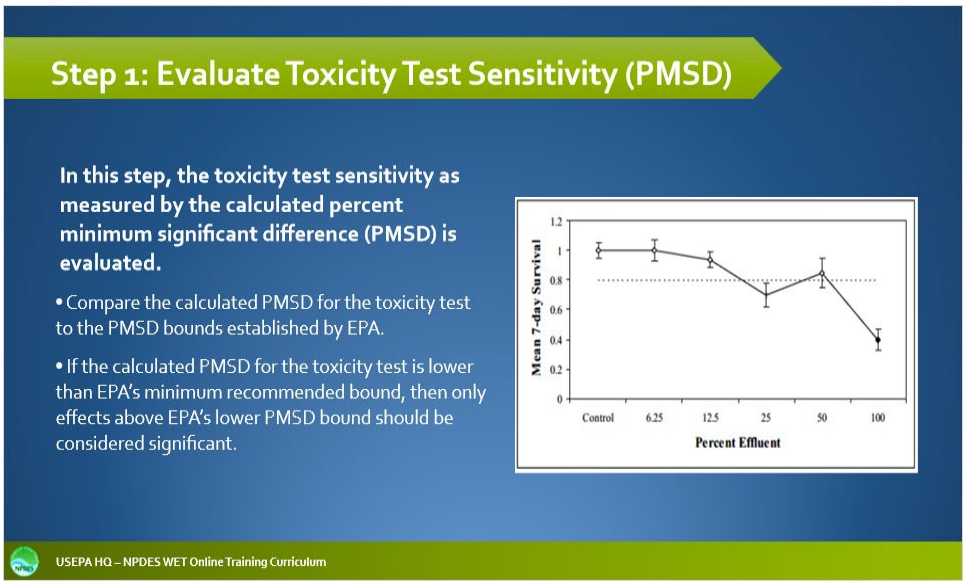
Notes:

The next review step is to evaluate within-treatment variability for the treatment identified as significantly different from the control. If the variability, measured by the calculated standard deviation or coefficient of variation, for that treatment is considerably greater than for the other treatments, then the response of the individual replicates should be reviewed. This review may show that a single outlier replicate has biased the treatment mean. If this is the case and the responses from all but the single outlier replicate are consistent with the control response, EPA recommends that the toxicity test be conducted again with a newly collected sample.



Notes:

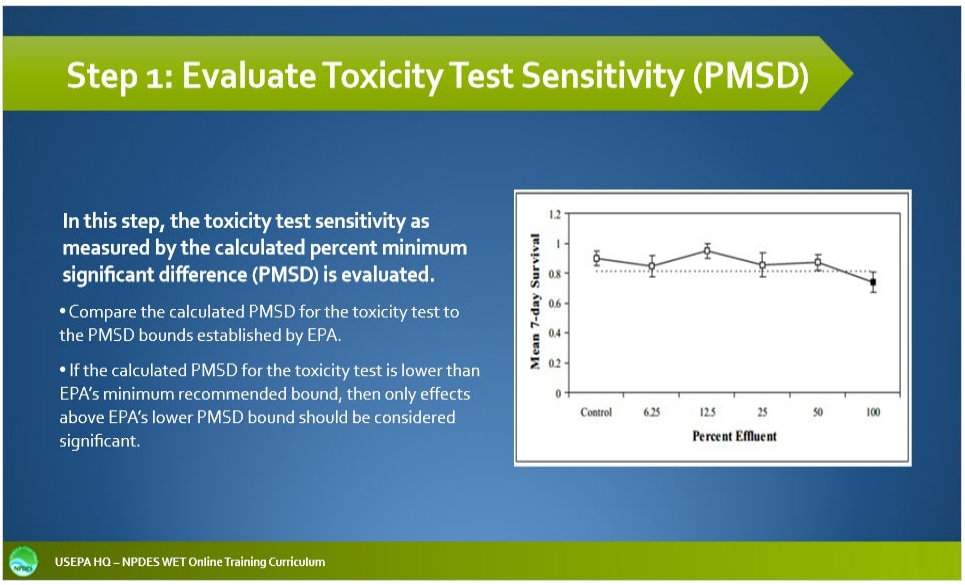
The final review step is to evaluate the toxicity test sensitivity using the percent minimum significant difference, or PMSD. If the toxicity test sensitivity is low, or a high PMSD value is associated with the test data, large effects at higher test concentrations may not be detected as statistically significant. EPA recommends collecting a new sample or samples and conducting a new toxicity test. If the toxicity test sensitivity is moderate to high, or a low PMSD value is associated with the test data, it is likely that the significantly different treatment is the result of a Type I error. A Type I error is the error of incorrectly rejecting the null hypothesis or assuming that the treatment is significantly different from the control, when in fact the null hypothesis is true, the treatment is not significantly different from the control. In this situation, the intermediate test concentration that was determined by hypothesis testing to be statistically different from the control should be considered anomalous, and the NOEC should be determined as the highest test concentration that was not significantly different from the control. The test results should still note that the 25 percent test concentration was statistically different from the control but was considered anomalous due to analysis of the CRR and the recommended review steps. Therefore, EPA recommends that the toxicity test be conducted again with a newly collected sample.



Notes:

For this CRR, it is important to evaluate test sensitivity by comparing the test percent minimum significant difference, or PMSD, to minimum and maximum PMSD bounds recommended by EPA and discussed previously in this module. If the PMSD is lower than the minimum PMSD bound, only effects larger than the minimum PMSD bound should be considered significant. For example, if the minimum PMSD for a method is 15 percent and the calculated test PMSD is 10 percent, only effects greater than 15 percent difference compared to the control should be considered as significantly different as noted in previous slides. If test sensitivity is low, where the test PMSD is above the maximum PMSD bound, EPA recommends that the toxicity test be conducted again with a newly collected sample.

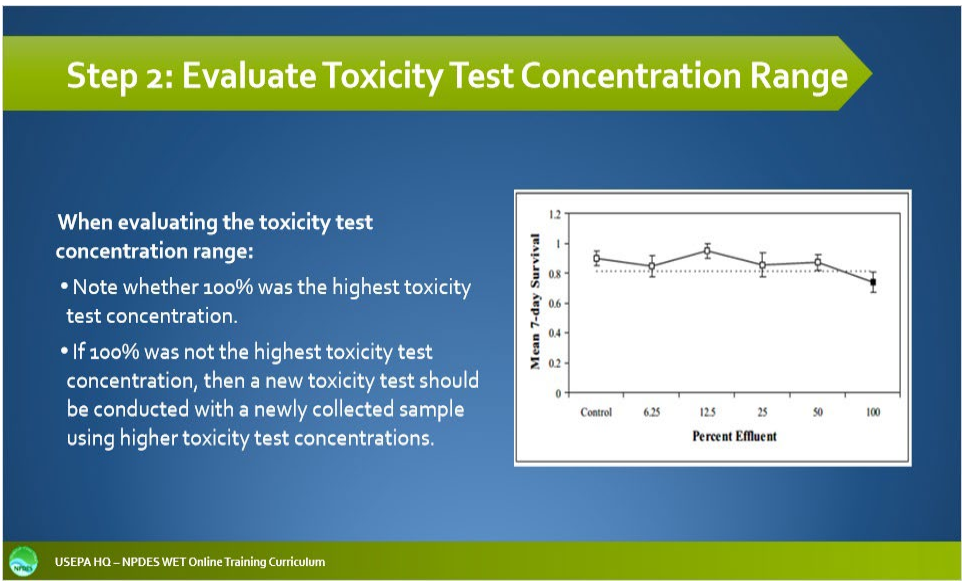
If the toxicity test sensitivity is moderate, where the test PMSD is within minimum and maximum PMSD bound, the test results should be considered valid, and the NOEC should be reported as the concentration below the LOEC. For the case depicted in the graph on this slide, a NOEC of 12.5 percent should be reported.



Notes:

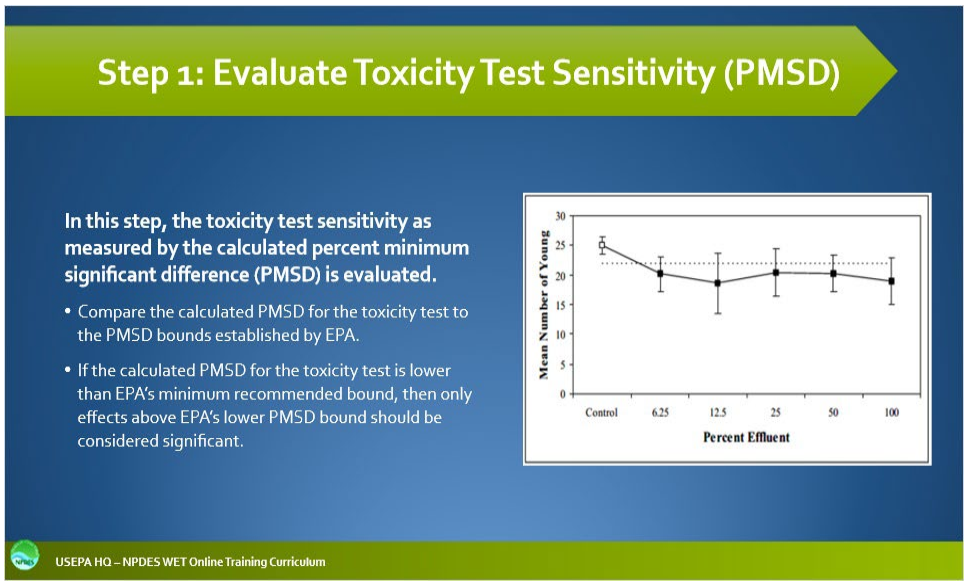
For this CRR, it is also important to evaluate toxicity test sensitivity by comparing the test percent minimum significant difference, or PMSD, to minimum and maximum PMSD bounds recommended by EPA and discussed previously in this module. If the PMSD is lower than the minimum PMSD bound, only effects larger than the minimum PMSD bound should be considered significant. For example, if the minimum PMSD for a method is 15 percent and the calculated test PMSD is 10 percent, only effects greater than 15 percent difference compared to the control should be considered significant. If test sensitivity is low, where the test PMSD is above the maximum PMSD, EPA recommends that the toxicity test be conducted again with a newly collected sample.

If the toxicity test sensitivity is moderate, where the test PMSD is within minimum and maximum PMSD bounds, the test results should be considered valid, and the NOEC should be reported as the concentration below the LOEC. For the case depicted in the graph on this slide, a NOEC of 50 percent should be reported.



Notes:

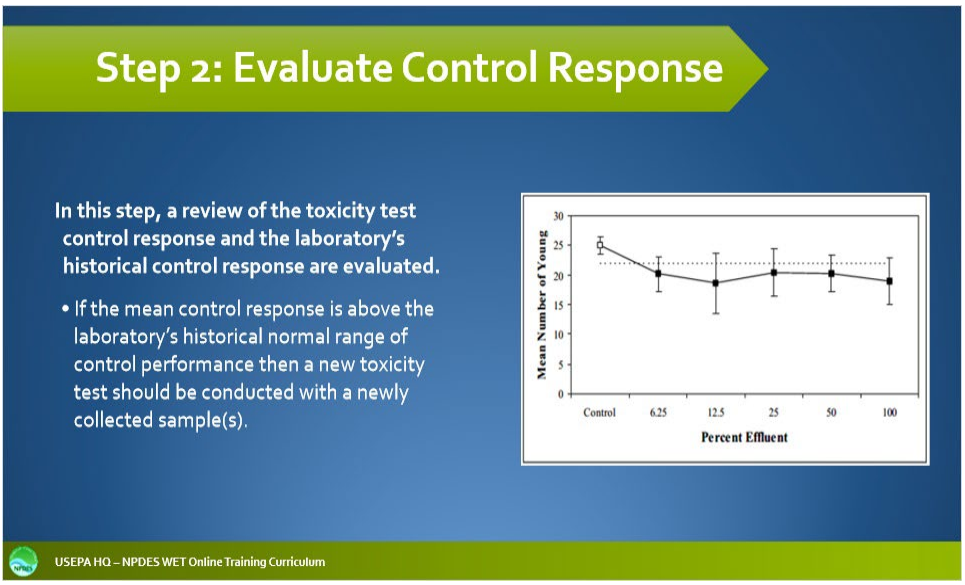
When this CRR occurs, the concentrations used for toxicity testing should be evaluated in future toxicity tests using this effluent. If the highest toxicity test concentration used in the toxicity test was less than 100 percent, future toxicity testing using this effluent should include at least one higher test concentration. If this response relationship occurs commonly with a given effluent, future toxicity testing of the effluent should use a dilution factor of greater than 0.5 such that test concentrations closer to the 100 percent effluent test concentration are used, for example a dilution factor of 0.65 would provide a test concentration series of 18, 27, 42, 65 and 100 percent effluent. This would provide a better opportunity to confirm effects that may exist at the upper end of the test concentration range. This approach should be used only if historical testing of the effluent indicates consistency, and the effect test concentration is not likely to fall below the adjusted test concentration series.



Notes:

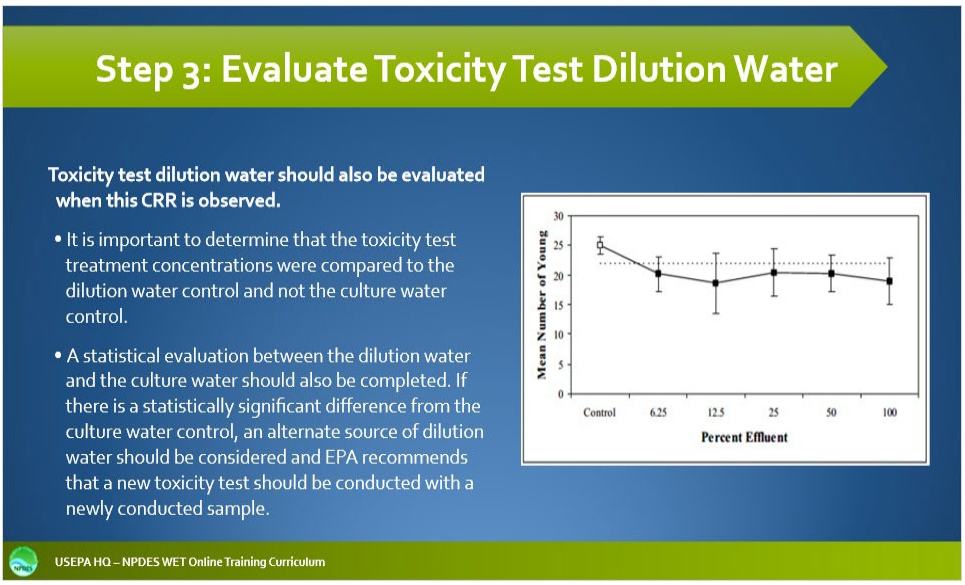
For this CRR, it is also important to evaluate toxicity test sensitivity by comparing the test percent minimum significant difference, or PMSD, to minimum and maximum PMSD bounds recommended by EPA and discussed previously in this module. If the PMSD is lower than the minimum PMSD bound, only effects larger than the minimum PMSD bound should be considered significant. For example, if the minimum PMSD for a method is 15 percent and the calculated test PMSD is 10 percent, only effects greater than 15 percent difference compared to the control should be considered significant. If test sensitivity is low, where the test PMSD is above the maximum PMSD, EPA recommends that the toxicity test be conducted again with a newly collected sample.

If the toxicity test sensitivity is moderate, where the test PMSD is within minimum and maximum PMSD bounds, the test results should be considered valid, and the NOEC should be reported as the concentration below the LOEC. For the case depicted in the graph on this slide, a NOEC of 50 percent should be reported.



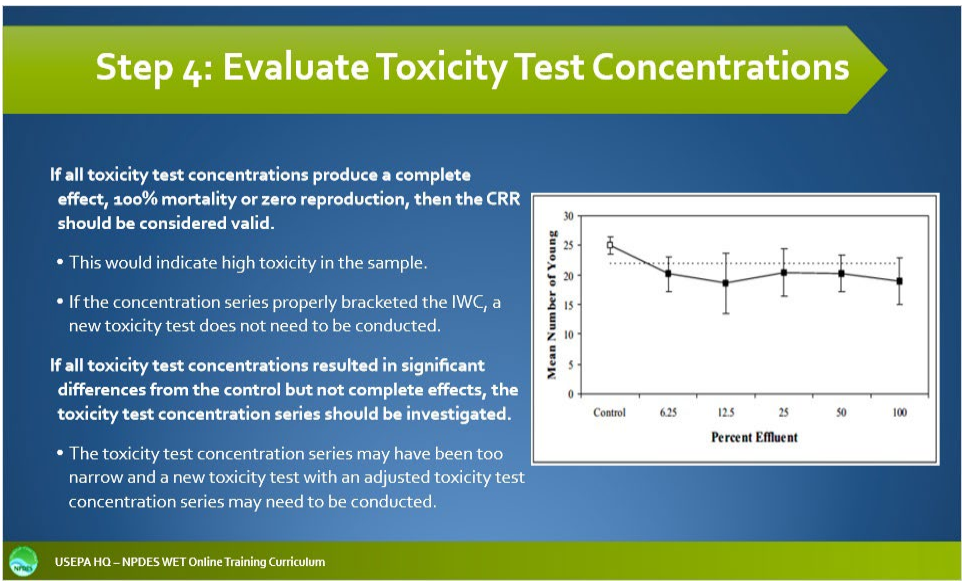
Notes:

Additionally, this CRR could result from an unusually high response in the control treatment. Laboratories are encouraged to track the performance of control results in toxicity tests conducted over time. When the response relationship depicted in this slide is encountered, the control response for the test should be compared to the historic control performance in the laboratory using the given dilution water. If the mean control response is above the normal range for that laboratory and dilution water, EPA recommends that the toxicity test be conducted again with a newly collected sample.



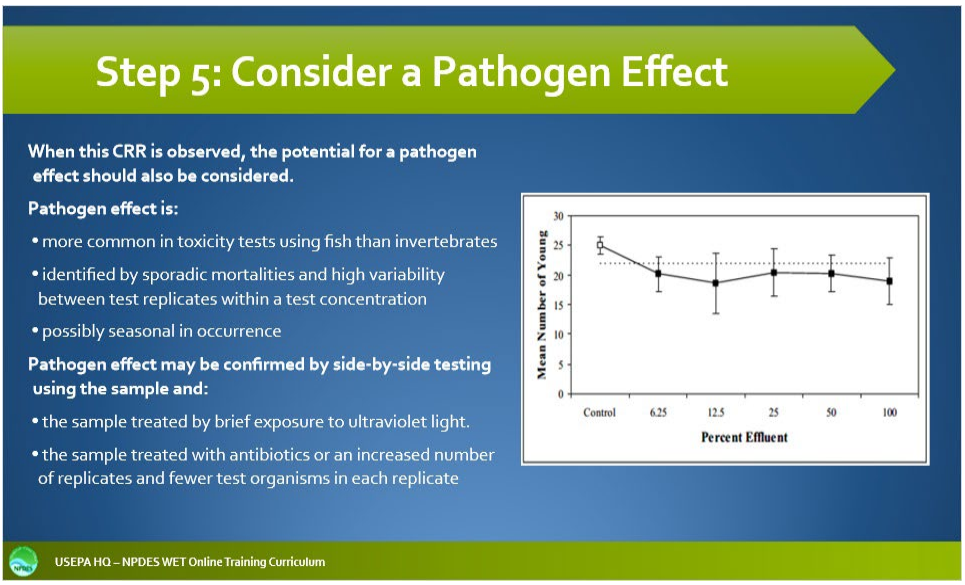
Notes:

The dilution water used should also be evaluated. The improper use of dilution waters and dilution water controls could cause this type of CRR to be encountered. It should be confirmed that the toxicity test treatment concentrations were compared to the dilution water control and not a culture water control. A statistical comparison of the dilution water control and the culture water control should also be made if they are from different sources. If the dilution water control shows a statistically significant difference from the culture water control, alternate dilution waters should be considered, and EPA recommends that the toxicity test be conducted again with a newly collected sample.



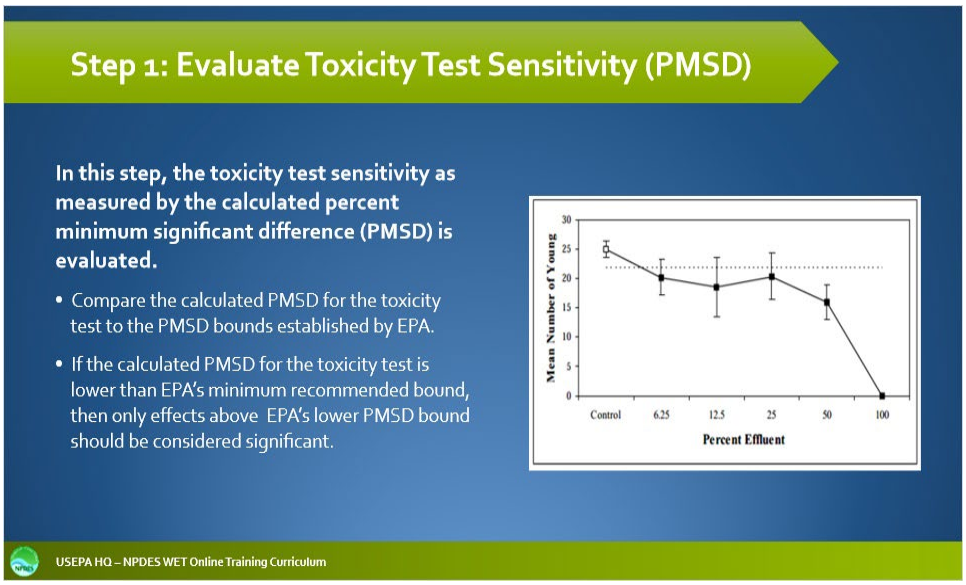
Notes:

If all toxicity test concentrations produce a complete effect, for example 100 percent mortality or zero reproduction, a flat CRR will result. This CRR should be considered valid, and it indicates high toxicity in the sample. Assuming the test concentration range used in the toxicity test brackets the permitted IWC, it is not necessary to conduct the toxicity test again, since the test result clearly indicates toxicity. If all test concentrations were significantly different from the control but did not produce complete effects as depicted in this graph, the test concentration series should be investigated. It is possible that the test concentration range used for the toxicity test was too narrow to distinguish a shallow sloped concentration-response relationship. Test concentrations may not have been low enough to produce no significant effect and may not have been high enough to produce severe effects. If this situation is suspected, EPA recommends that the toxicity test be conducted again with a newly collected sample using an expanded test concentration series range. Test concentrations that are lower than those used in the previous toxicity test should be added. Test concentrations that are higher than those used in the previous toxicity test should also be added, if possible, to assist in determining a CRR.



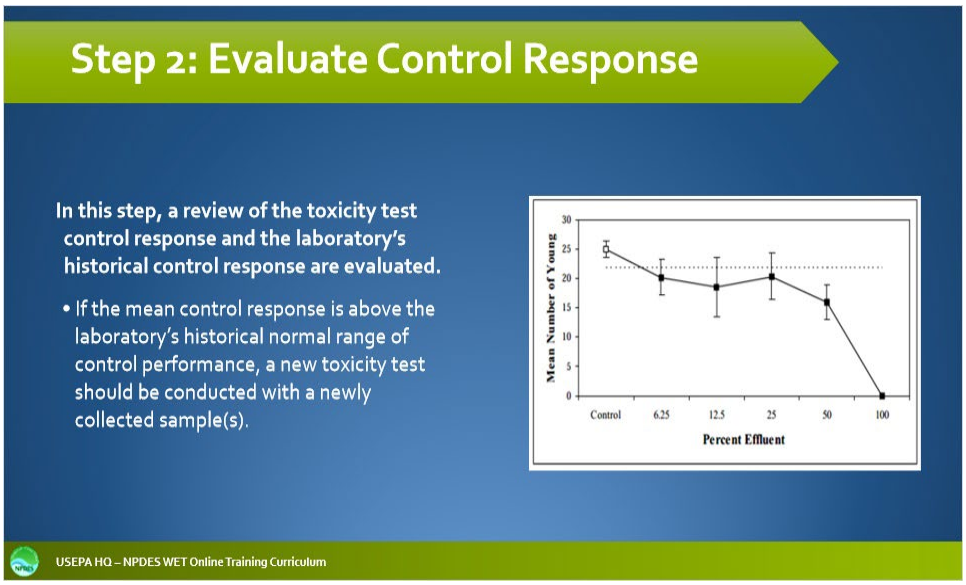
Notes:

The concentration-response pattern depicted here could also be due to the presence of pathogens in the sample. The most common identifier of pathogen effects are sporadic mortalities and extremely high variability between the toxicity test replicates. The pathogen effect is more common in toxicity tests using fish species than in invertebrate testing. This pathogen effect also may be evident only in chronic tests and not in acute tests. Pathogen effects also may be seasonal in occurrence. If the within-treatment coefficient of variation, or CVs, for survival are greater than 40 percent for test concentrations and relatively small for control replicates in standard synthetic water, pathogen effects should be considered. If pathogen effects are suspected in the sample, this may be confirmed in subsequent side-by-side toxicity testing using the sample and the sample treated by brief exposure to ultraviolet, or UV, light, the addition of antibiotics or increasing the number of test replicates and using fewer test organisms in each test replicate. If pathogen effects in the sample are confirmed, EPA recommends that toxicity test be conducted again with a newly collected sample and the regulatory authority should be consulted prior to changing toxicity testing procedures.



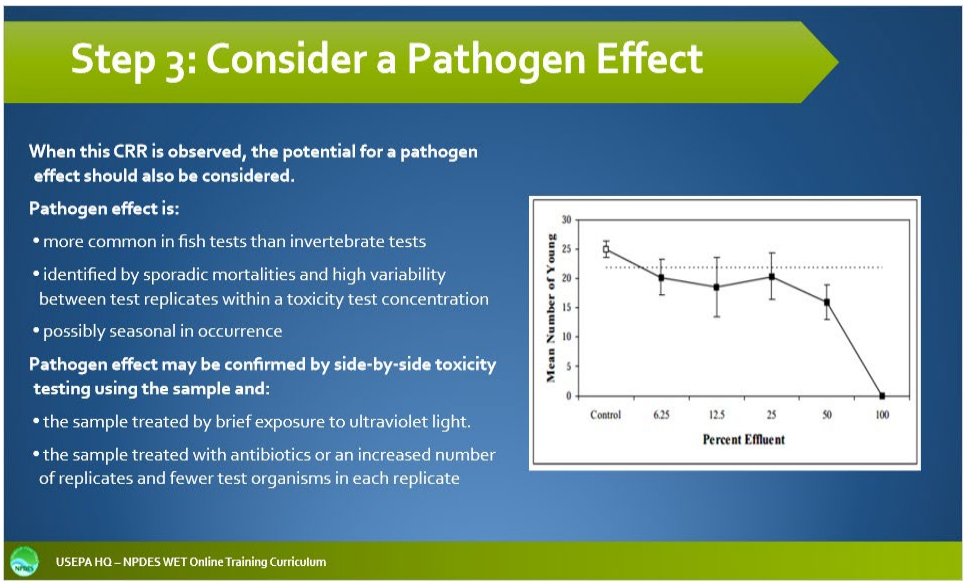
Notes:

For this CRR, it is also important to evaluate toxicity test sensitivity by comparing the test percent minimum significant difference, or PMSD, to minimum and maximum PMSD bounds recommended by EPA and discussed previously in this module. If the PMSD is lower than the minimum PMSD bound, only effects larger than the minimum PMSD bound should be considered significant. For example, if the minimum PMSD for a method is 15 percent and the calculated test PMSD is 10 percent, only effects greater than 15 percent difference compared to the control should be considered significant. For the case depicted in the graph on this slide, a NOEC of 50 percent should be reported.



Notes:

Additionally, this CRR could result from an unusually high response in the control treatment. Laboratories are encouraged to track the performance of control results in toxicity tests conducted over time. When the response pattern depicted in this slide is encountered, the control response for the toxicity test should be compared to the historic control performance in the laboratory using the given dilution water. If the mean control response is above the normal range for that laboratory and dilution water, EPA recommends that the toxicity test be conducted again with a newly collected sample.



Notes:

The concentration-response pattern depicted here could also be due to the presence of pathogens in the sample. The most common identifier of pathogen effects are sporadic mortalities and extremely high variability between the toxicity test replicates. The pathogen effect is more common in toxicity tests using fish species than in invertebrate testing. This pathogen effect may also be evident only in chronic tests and not in acute tests. Pathogen effects also may be seasonal in occurrence. If the within-treatment coefficient of variation, or CVs, for survival are greater than 40 percent for test concentrations and relatively small for control replicates in standard synthetic water, pathogen effects should be considered. If pathogen effects are suspected in the sample, this may be confirmed in subsequent side-by-side toxicity testing using the sample and the sample treated by brief exposure to ultraviolet, or UV, light, the addition of antibiotics or increasing the number of test replicates and using fewer test organisms in each test replicate. If pathogen effects in the sample are confirmed, EPA recommends that the toxicity test be conducted again with a newly collected sample and the regulatory authority should be consulted prior to changing toxicity testing procedures.