Responses to Public Comments on the Office of Pesticide Program's Science Policy Documents:

Assigning Values to Nondetected/Nonquantified Pesticide Residues in Human Health Dietary Exposure Assessments (EPA Docket #OPP-00570)

and

A Statistical Method for Incorporating Non-detected Pesticide Residues in Human Health Dietary Exposure Assessments (EPA Docket #OPP-00571)

Health Effects Division
Office of Pesticide Programs
US Environmental Protection Agency
Washington, DC 20460

March 23, 2000

List of Acronyms

aPAD	acute Population Adjusted Dose
IR-4	Interregional Project #4
LLMV	Lower Limit of Method Validation
LOD	Limit of Detection
LOQ	Limit of Quantitation
ND	Non-detect
OPP	Office of Pesticide Programs
PDP	Pesticide Data Program

Table of Contents

I. Introduction

A. Background

On December 4, 1998 the US Environmental Protection Agency (USEPA), Office of Pesticide Programs (OPP) issued in the *Federal Register* a Notice of Availability (along with a request for comment) regarding two science policy paper on the incorporation of "non-detectable" pesticide residues into acute and chronic risk assessments. These documents, entitled *Assigning Values To Nondetected/Nonquantified Pesticide Residues into Human Health Dietary Exposure Assessments* and *A Statistical Method for Incorporating Nondetected Pesticide Residues into Human Health Dietary Exposure Assessments*, discussed a set of issues dealing with the selection of an appropriate residue concentration value to use in risk assessments when a treated agricultural commodity contains pesticide residues at a concentration of less than the analytical detection limit.

Many parties commented on the science policy issues (submitted under dockets OPP-00570 and OPP-00571). They included pesticide registrants, environmental and public interest groups, and consultants. All comments and recommendations were reviewed by OPP and incorporated as appropriate, into the current revised science policy document. The comments ranged in specificity. Some commenters addressed the general policy and its rationale as well as all of the specific questions posed, while other reviewers provided detailed comments only on certain aspects of the proposed policy. A listing of the names and affiliations of the parties submitting comments appears at the end of this document (See Section IV– List of Commenters)

B. Organization of this Document

This response package contains OPP's combined responses to the comments raised in the two science policy papers. In an effort to address the overall issue of non-detectable residues (and OPP's recommended approach to incorporating these non-detectable residues into its risk assessments), the two policy documents have been combined into one entitled *Assigning Values to Non-detected/Non-quantified Pesticide Residues in Human Health Dietary Exposure Assessments*. The document is organized by topic area, each of which contains a summary of the key elements of the 1998 science policy guidance, a synopsis of the public comments which were submitted, and the Agency's response. These responses include OPP discussion of the comments received on the questions posed by OPP in each of the following two science policy papers:

> *"Assigning Values to Nondetectable Pesticide Residues in Human Health Dietary Exposure Assessments*"
>
> 1. Under what circumstances would either ½ LOD or LOQ for NDs significantly underestimate or overestimate pesticide food residue exposure? Does any available information demonstrate that this method either underestimates or overestimates pesticide food residue exposure?

2. Should EPA consider a different approach for incorporating nondetectable samples into risk assessments depending on the type of risk assessment being performed (i.e., chronic risks, acute risks, short-term risks (Section 18's))?

3. Are the methods for determining LOD and LOQ adequately defined?

4. Would this policy have any implications for international trade?

"A Statistical Method for Incorporating Nondetectable Pesticide Residues into Human Health Dietary Exposure Assessments"

1. Are other methods available which may be preferable to the methods described in this paper for statistically estimating the distribution or mean values of nondetectable residue samples?

2. Under what circumstances, if any, would use of Cohen's method not be considered reliable or appropriate?

To organize the responses to the comments received on these questions, OPP has combined them into several larger topic areas under each of the original policy documents:

"Assigning Values to Nondetectable Pesticide Residues in Human Health Dietary Exposure Assessments"

- ❖ Risk Management vs Policy
- ❖ What Value to Use for Non-detected or Non-quantified Residues
- ❖ Sensitivity Analysis
- ❖ Guidance on Operational Determination of LOD and LOQ
- ❖ Additional Refinements, Suggestions, and Clarifications

"A Statistical Method for Incorporating Nondetectable Pesticide Residues into Human Health Dietary Exposure Assessments"

- ❖ Recommendations for Application of Methods
- ❖ Other Available Methods
- ❖ Suggested Clarifications/Corrections

A brief summary of the comments in each topic area is provided immediately prior to the detailed responses in the relevant section.

II. Response to Comments on LOD/LOQ Issues

A. Risk Management vs. Policy

✔Science Policy vs. Risk Management vs. Risk Assessment

Comment: One commenter believed that the policies were mislabeled and contended that they were not science policies, but rather risk management policies which are science-based. Risk management policies should incorporate additional views based on the current state of knowledge and would be based on data collected in a scientific (objective) manner. The commenter recommended that future determinations of pesticide exposure rely extensively on monitoring of actual exposure under conditions of real-world use.

OPP Response: OPP disagrees with the commenter. The original intent of the document was to propose a method to consistently deal with residues that are non-detectable or non-quantifiable for use in conducting risk and exposure assessments for pesticides. The science policy document was issued to provide background information on this issue and to propose a policy for appropriately incorporating ND values into OPP risk assessments. It provides support for the proposed policy's reasonableness and validity, asks for comments on why this policy should or should not be considered, and seeks suggested refinements and improvements. OPP believes that the proposed policy falls more into the realm of "risk assessment/risk characterization" than "risk management" in that it is designed to permit the risk assessor to develop an estimate of pesticide exposure and characterize the uncertainties in this estimate for appropriate consideration by the risk manager.

B. What Value to Use for Non-detected/Non-quantified Residues

✔What to Use for Non-detected Residues
✔What to Use for Non-quantified Residues
✔What to Use for Residue Measurements Between LOD and LOQ
✔What to Use for Untreated Commodity

Comment: Several commenters supported the use of ½ LOD as a default value when residues, if present, are below the LOD. One commenter emphasized that this should be seen as just a default, which should be able to be overridden with real data or additional information. The commenter indicated that if a crop or animal was not treated with the chemical and there was also no use of it in or near the production area in previous years and the chemical does not have mobility characteristics that might bring it in from somewhere else, then "there could truly be zero amount present, and zero might be the best value to plug in." Another commenter, also

supporting the use of ½ LOD or ½ LOQ, indicated that values of zero are appropriate for that proportion of the residue data set corresponding to the percentage of the commodity known not to be treated. This commenter urged OPP to use a value less than ½ the LOD or ½ the LOQ for those situations in which no detectable residues are observed in trials conducted at exaggerated application rates. Another commenter supported the use of true zero in cases where residues would clearly not be present on the food item as it enters interstate commerce, and the use of ½ LOD or ½ LOQ is appropriate to use as a central tendency for establishing a more accurate estimate of residues present on the food item regardless of the type of risk assessment. Several commenters encouraged the Agency to prefer that the analyst's measured value (if that measured value is between the LOD and LOQ) in preference to the selection of a default value of ½ LOD or ½ LOQ.

Using a similar line of reasoning, a commenter indicated that for certain use patterns (e.g., seed treatments and applications to dormant fruit and nut trees), OPP should assign a value of zero to residues falling below the LOQ and that assigning a value of ½ the LOD or ½ the LOQ would significantly overestimate pesticide food residue exposure when applied to seed treatments at low application rates. Another commenter indicated that EPA should develop a reasonable expectation about the residues in ND samples, using information independent of the analytical sampling. The commenter stated that EPA should assess whether the ND samples arise from circumstances where the pesticide in question probably was not used (in which case, the commenter claims zero is reasonable) regardless of whether the NDs arose from circumstances where the survey showed some values below detection limits (in which case EPA may want to use one-half the detection limit as a crude estimator).

Another commenter, voicing a similar opinion, believed that adopting a policy of assigning ½ LOD for non-detectable residues for purposes of calculating possible or theoretical risk appears to be in conflict with the concept of "essentially zero" exposure (and "essentially zero" risk) which had also been proposed for separate comment. That policy, the commenter continued, rests on the premise that some exposures (and risks) are just to small to be of concern. The commenter stated:

> We agree with the Agency on this premise. It seems to us that residues which are so small as to not warrant regulation should not be assigned a finite value, but should be treated as true zeroes for purposes of risk assessment. We recognize that in some cases, e.g., where data sets include a significant number of measurable residue values, assigning true zeros to all non-detectable residues would likely underestimate risk and therefore is probably not appropriate. However, for some data sets which consist entirely, or almost entirely, of non-detects, assigning ½ LOD to all non-detects may considerably overestimate risk.

OPP Response: OPP is in basic agreement with the commenters. OPP notes that a "true zero" value is routinely incorporated into assessments for that portion of the crop which is not treated. Also, ½ LOD or ½ LOQ is used as a "starting point" in OPP's assessments, which can always be re-evaluated by the OPP risk assessor as necessary. A "sensitivity analysis" is regularly done in those situations where the assessor believes that the ½ LOD or ½ LOQ assignment to

ND's is significantly affecting the risk and exposure estimate or the risk management decision. In all cases where this is believed to be occurring, the risk assessor will attempt to produce a better estimate for ND values and will fully inform the risk manager as part of the risk characterization. In those cases where the risk assessor thinks that a zero or near zero estimate for residue concentrations in agricultural commodities is a truer representation of actual residues than the ½ LOD or ½ LOQ default, the assessor is free to insert this value into the risk assessment, provided adequate justification is given. In fact, this option has been explicitly discussed in HED SOP 99.6, has been summarized in the revised policy document, and is excerpted, in part, below (US EPA, 1999):

> ... [It may be appropriate in certain cases to judge that the ND values from the monitoring data are "essentially zero," particularly if a substantial portion of the measured residue values are less than the analytical detection limit (and would therefore ordinarily be replaced by ½ LOD). In these instances, it may be appropriate to introduce a value of zero ppm (or near zero) as a residue value (in place of ½ LOD) for the ND measurements in the risk assessment. This judgement should be made on a case-by-case basis, with the reviewer bringing a wide range of information to bear on proper valuation of the NDs, including the nature of distribution of the values above the detection limit, the percent of the crop which is treated, and information on the processing of commodities before sampling;

Comment: One commenter asked at what point the LOD or LOQ information became irrelevant. He cited an example in which detectable residues were found in only 10% of the analyzed samples and asked if 90% of the non-detects should be assigned ½ the LOD or ½ the LOQ values.

In addition, the commenter raised a concern about what should be done for chemicals for which the LOD is only a small fraction of the established tolerance (e.g., 1/100) and whether these chemicals be treated the same or differently as chemicals for which the LOD is, for example, a larger fraction of the tolerance (e.g., 1/10). These risk management decisions, the commenter continued, need to be made to promote consistency in the application of the regulations and also allow flexibility to address unusual situations and ignore the absurd.

OPP Response: In the example cited by the commenter in which detectable residues were found in only 10% of the analyzed samples, routine OPP practice would be to use "true zeroes" for that portion of the crop which was not treated. Thus (in an example consistent with the example cited by the commenter), if it was estimated that 20% of the crop was treated and only 10% of the analyzed samples contained residues, OPP would assign "true zero" for the 80% of the crop which was not treated, ½ LOD or ½ LOQ for the 10% of the crop which was treated but contained non-detectable or nonquantifiable residues, and assign the remaining 10% of the values the actual measured analytical concentrations. As described in the response to the previous comment, the sensitivity of the exposure and risk estimate to the assignment of ½ LOD (or ½ LOQ) value to the treated non-detects would be assessed and, if the effect was significant, the risk assessor would re-evaluate the ½ LOD or ½ LOQ value assigned to the treated detects and be free to change this assignment in accordance with the circumstances of the specific situation.

With respect to the commenter's question about those situations in which the LOD is only a small fraction of the established tolerance and whether these chemicals should be treated the same or differently as chemicals for which the LOD is, for example, a larger fraction of the tolerance, OPP is unsure of the point the commenter is attempting to make. As an example, the commenter asks about a hypothetical situation in which, for one crop, the LOD is a small fraction of a tolerance (e.g., perhaps the tolerance is 100 ppm and the LOD is 1/100 of that or 1 ppm) and, for another crop, the LOD is a large fraction of the tolerance (e.g., perhaps the tolerance is 0.01 ppm and the LOD is one-tenth of that or 0.001 ppm) and wonders if these would be treated the same or differently by OPP for risk assessment purposes.

OPP sees no reason why this situation should necessitate a change in the assignment procedure for LODs or LOQs or why the procedure for assigning default values should be adjusted depending on whether the ratio is big or small. The tolerance represents the maximum legal residue which may be present on an agricultural commodity as it enters interstate commerce and generally reflects the highest residue found following field trials in which the maximum application rate is applied to a crop which is harvested at the shortest pre-harvest interval permitted on the label. The LOD is an operationally defined quantity which reflects the lowest concentration of an analyte in a sample at which one can say with reasonable certainty (e.g., 99%) that the analyte is present. OPP sees no connection between, and has no intention of linking, these two disparate values. The value estimated for an LOD or LOQ is in no way linked to the residue expected on the crops following maximum application rate or minimum pre-harvest interval.

Comment: Several commenters generally agreed with and supported EPA's current policy for assigning values to non-detected residues. As stated by one:

> Decades of federal experience with pesticide regulation have shown that a nondetection does not necessarily mean that residues are not present. As detection technologies have become more sensitive, foods once thought to contain zero residues actually contain small amounts of residues that were previously undetectable. The Delaney zero-cancer risk standard for processed food did not become bothersome for chemical companies and agricultural growers until detection technologies began picking up low but real pesticide residues on food. The clear lesson is that zero does not necessarily mean zero.

The commenter stated that EPA should continue assigning ½ LOD or ½ LOQ to residue samples below the LOD or LOQ associated with that portion of the crop believed to have been treated with the pesticide for *chronic* risks. This, the commenter asserts, is scientifically sound and reasonably health protective. Nevertheless, several commenters expressed concern over EPA's decision to discontinue the use of the full LOD or LOQ for *acute* risk assessments. For acute risk assessments, it is the actual residue, one commenter claims, and not the mean residue, that is of concern. Stating that the full LOD represents a reasonable upper bound for non-detected residues for acute exposure assessments, one commenter stated that EPA should continue using this value. EPA' s only justification for selecting the ½ LOD or ½ LOQ value for use in risk assessments, the commenter claims, is that it will make the policy clear to the public.

"While maintaining clear and consistent pesticide policies," the commenter continues, " is a laudable goal, a health protective approach should not be abandoned simply to make a policy appear more consistent." Sharing this concern, another commenter (citing the language of Title IV, Section 405 of the FQPA) stated that the language of the law implied that it was appropriate for the Agency to consider percent crop treated information "for the purpose of assessing chronic dietary risk." However, the same commenter also cited the accompanying report language in HR Report 104-669, Part 1 in which such a distinction between acute and chronic risk in the context of using data on the percent of crop treated is not created. The commenter requested that the EPA clarify its position on this issue and urged the Agency to adhere to the strict interpretation of "reasonable certainty of no harm," especially for children and other high risk populations when doing so.

OPP Response: OPP understands that the commenters believe that while it is appropriate to use ½ LOD or ½ LOQ to represent non-detected residues in treated commodities in the case of chronic exposure assessments, the commenters feel it is more appropriate to use the full LOD (or LOQ) for acute assessments. This is how ND values were incorporated into the risk assessment prior to use of probabilistic assessments. With the goal of probabilistic risk assessment being to provide more accurate estimates of exposure and risk while at the same time avoiding underestimates of risk, OPP believes that it is appropriate to produce better estimates of residue values associated with non-detectable residues and we believe that the use of ½ LOD (or ½ LOQ) is a more appropriate estimate to use. In any case, in those situations where OPP believes that this use of ½ LOD or ½ LOQ may underestimate risks, the risk assessor will perform a sensitivity analysis to verify that the use of ½ LOD or ½ LOQ has no substantial effect on either the exposure/risk estimate or the resulting risk management decision.

With respect to the commenter's concern about the legal permissibility of incorporation of percent crop treated into acute risk assessments, OPP notes that the FQPA amendments to FFDCA do not prohibit the use of any type of information in the performance of risk assessments. Although the FQPA does contain conditions on the use of percent crop data when OPP performs chronic risk from pesticides in food, it is silent with respect to whether those conditions apply to risk assessments involving acute exposures to pesticides in food.

The Agency has decided to exercise its discretion to use available percent crop treated information in its acute exposure risk assessments, when the data are judged to be reliable. Using this information will produce a more realistic estimate of potential exposure, and thus a sounder basis on which to make regulatory decisions. Although the conditions applicable to chronic food risk assessments do not automatically apply, OPP will use its judgment to reexamine periodically those acute pesticide food residue risk assessments whenever the percent crop treated information played a significant role in the risk assessment and there is any reason to believe that the percent crop treated may have increased substantially above the value used in the assessment.

Comment: One commenter stated that he agreed that LOQ and LOD should be clearly demonstrated and that acceptable procedural recoveries from each matrix/chemical/

instrument combination at the LOQ should be obtained. Several commenters stated that it would be more appropriate to use one-half the Lower Limit of Method Validation, or ½ LLMV (in those cases where it is properly determined), rather than the full LLMV. One commenter went on to state that the Science Policy document indicates that the LLMV must be used if the LOQ and/or LOD are not properly determined or estimated. The commenter continued:

> If the LLMV is validated properly (acceptable procedural recoveries), values below the LLMV should be entered at ½ LLMV as long as EPA performance criteria are met (mean recoveries fall between 70-120% and the relative standard deviation is less than or equal to 20%). Since OPP has sufficient confidence in the LLMV to set tolerances and because the LLMV represents the lower limit of reliable quantitation (in the presence of substrate), the Agency should be comfortable with the use of ½ LLMV for non-detected residues if empirical performance criteria are met.

OPP Response: OPP greatly prefers that registrants or other data submitters estimate an LOD or LOQ using rigorous statistical methods. When this is done, OPP is willing to assume (as an initial assumption subject to change) that ND residues in treated commodities are present at ½ the LOD or ½ the LOQ. When, instead of determining an LOD or LOQ in accordance with any of a number of acceptable and widespread conventions, the data submitter simply determines a LLMV, the statistical rigor is absent (if an estimate of the LOD or LOQ cannot subsequently be obtained from the data). The reason OPP is proposing that the full LLMV be used in a risk assessment when an LOD or LOQ is not properly determined should be readily apparent: OPP does not want to create a set of circumstances in which is it advantageous to use a poorly documented non-statistical procedure to estimate a default value for use in risk assessment which might be lower than would be produced with comparable data using a rigorous procedure. As an example, suppose that the LLMV is 0.1 ppm and three recoveries studies are performed yielding 70% recovery, 72% recovery, and 120% recovery (i.e., spike at 0.1 ppm and obtain 0.070 ppm, 0.072 ppm, and 0.120 ppm). These three recovery percentages are all acceptable (and average close to 87%) as per OPP test guidelines which recommend recoveries of between 70% and 120%. Using ½ LLMV, the residue used in a risk assessment would be ½ * 0.1 ppm or 0.05 ppm. Calculating the LOQ using a rigorous statistical procedure would produce an estimate of 0.1972 ppm for LOQ and an estimate of 0.0986 ppm for ½ LOQ. In this case, it would be more advantageous for the registrant or data submitters to conduct a poorly designed calibration study based on only 3 spike recovery samples and use the ½ LLMV of 0.05 ppm value than it would be to conduct a well designed study and use of the ½ LOQ value of 0.0986 for risk and exposure assessment purposes. OPP does not intend to create or introduce these kinds of incentives in its guidance and believes that if a LLMV study is reasonably designed, properly carried out, and rigorously analyzed, the registrant or data submitters should be capable of estimating a valid LOQ from the data (in which case OPP would use ½ LOQ in its estimates)

Comment: One comment expressed strong disagreement with using a weighted average of detection limits from treated commodities in which analytical laboratories detected no residues. Detection limits, the commenter reasoned, are highly specific to each pesticide, each method, and each laboratory, and may even drift with time. He stated that the use of an average detection limit

involving different methods, labs, and times has little, if any, scientific meaning.

The commenter stated that each time a laboratory obtains an analytical measurement of a sample below the detection limit, the Agency should require a notation of the sample's value as "ND" with a footnote stating the laboratory's current applicable detection limit for that pesticide. If OPP receives data tabulated in this way, the commenter continued, OPP should not immediately assign a value to the ND samples or insert numerical values into substitute data tabulations. Instead, OPP should "compare the data with zero assigned for all samples to the data with the footnoted detection limits assigned for all ND samples." The commenter stated that OPP policy is to replace the ND designation with zero in proportion to the percentage of the crop not treated and with an assigned value of ½ the detection limit for the remainder. The commenter indicates that this approach will likely fail since 1) OPP may not know the percentage of crops treated; 2) normal processing of foods may mix treated and untreated crops; 3) persons within or outside EPA will inevitably confuse tabulations with "ND" replaced by hypothetical estimates with accurate measures of residues; and 4) it may matter which specific ND samples get replaced with zero.

OPP Response: OPP has considered the above comments. OPP agrees that a method's LODs can vary from laboratory to laboratory and that it is important to account for the differences. OPP disagrees with the comments provided regarding the use of a weighted average. Although no specific alternative is suggested by the commenter, OPP infers that the approach would be to assign a ½ LOD value for each lab with the ½ LOD value determined by that laboratory's LOD. OPP maintains that the approach suggested by the commenter is substantially equivalent to the weighted average approach OPP is advocating and might differ (and then only slightly) only if there were differential assignment of samples to laboratories based on percent crop treated.

The commenter also expressed concern about not having reliable percent crop treated data and being unable to produce a reasonable estimate. If this information is not available, OPP assumes that 100% of the crop is treated. If this information is available, OPP assumes that untreated crops have zero residues. Also expressed by the commenter was a concern about normal processing of foods mixing treated and untreated crops. OPP has a specific policy on this (EPA, 1999) which states that if a commodity is assumed to be widely blended across a large region, percent crop treated is NOT a reliable surrogate for the probability of encountering a treated commodity and percent crop treated is not used. For example, if 10% of widely blended commodity such as corn is treated with a given pesticide, it by no means follows that only 10% of corn oil would potentially contain residues (it could conceivably be much greater). In contrast, if only 10% of apples are treated with the pesticide of interest, it is altogether reasonable to assume that only 10% of apples contain residues and that the remaining 90% of apples contain no pesticide. OPP's policy for dealing with percent crop treated does effectively consider and deal with these differences. Finally, with respect to the commenter's apparent concern that people "within or outside EPA will inevitably confuse tabulations with ND replaced by hypothetical estimates with accurate measures of residue," OPP believes that its guidance materials clearly

explain how OPP handles non-detected residues.

Comment: One commenter stated that the use of synthetic data sets that replace ND values with one-half of the detection limit can particularly upset the Agency's approach to modeling pesticide food residue exposure with Monte Carlo methods. The commenter advocated use of several existing software packages which work with left-censored data sets and states that misrepresent-ation of the distribution by using policy-generated values can distort parameterizations of the component populations.

OPP Response: This comment, while submitted in response to the LOD/LOQ science policy paper, is more appropriately discussed along with censored data. OPP responds to this comment in Section III of this response to comments.

Comment: One commenter indicated his belief that if the number of nondetects is relatively small (less than 10-15% of the total number of observations), assigning ½ the LOD to each ND is not a bad procedure. As the number of ND values become large (>15%), he continued, the estimates of quantities such as the arithmetic mean by simply substituting ½ the LOD or ½ the LOQ becomes problematic. The commenter recommended that in those situations in which more than 15% of the data is non-detect, that statistical estimation procedures designed to account for "censored" data be implemented.

OPP Response: OPP agrees that in most cases in which <10-15% of the total number of observations are LOD or LOQ, then assigning ½ LOD to each ND is a reasonable procedure. The commenter is suggesting that anytime >15% of the data are ND, then a statistical estimation procedure designed for "censored" data be implemented. OPP agrees that, in many instances, the use of statistical imputation procedure for censored data may produce a better estimate of <LOD residues than substitution methods with ½ LOD. However, OPP, as part of its risk characterization protocol, will evaluate when necessary the impact of the ½ LOD or ½ LOQ assumption and, if this is found to be significant, may decide to use statistical techniques designed for censored data. The risk characterization protocol states that if it is found by a sensitivity analysis that the designation of ND as ½ LOD (as opposed to zero or the full LOD) has a significant effect on the exposure estimate or risk management decision, then OPP will attempt to refine its risk estimates by using statistical estimation procedures designed to account for censored data.

C. Sensitivity Analysis

Comment: One commenter indicated that additional clarification would be useful in Section II.H of the original documents regarding the conduct of the sensitivity analysis. He stated that the section seems to enable, but not require, OPP to perform the sensitivity analysis by repeating the entire risk assessment (all commodities) with the specified alternative assumptions (either NDs all equal zero or all ND's are equal to ½ LOD or ½ LOQ). He stated that little discussion is provided as to whether the difference is significant or how the results of this effort

will be used.

OPP Response: The commenter is correct in that OPP does not plan to routinely repeat the risk assessment with the specified alternative assumptions, but may choose to do so where circumstances warrant. This will generally be judged on a case-by-case basis by the risk assessor.

If the difference in exposure estimates is significant (i.e., the risk or exposure estimate or risk management decision changes when the alternative assumption of either all ND's are zero or all ND's are the full LOD or LOQ is used), this information will be carefully considered and conveyed by the risk assessor to the risk manager. The risk assessor will attempt to refine his or her estimate of the low-end residue concentrations (perhaps using statistical techniques for censored data), recognizing that the assumption that all treated non-detects contain residues of ½ LOD (or ½ LOQ) is a significant factor in determining the outcome of the risk assessment. If adequate data are unavailable for this refinement, the risk manager may request that the registrant or data submitters provide additional information.

Comment: Another commenter disagreed with OPP that a sensitivity analysis is a supplemental procedure to do only under "certain" circumstances (namely, when the Agency believes that substitution of zero for one-half the detection limit significantly affects the risk assessment or regulatory decision). The commenter stated that EPA will not know when this situation arises without conducting a sensitivity analysis first.

OPP Response: This topic was considered in the previous response. To date, OPP has found through the numerous sensitivity analyses that the ½ LOD /LOQ assignment has had no significant effect on the risk and exposure estimates at the 99.9th percentile of estimated acute exposure to residues in food. OPP believes that the risk assessor is in the best position to judge whether a sensitivity analysis is a necessary and useful part of the risk characterization and intends to leave the determination of the need for a sensitivity analysis to the discretion of risk assessor. Obviously, if a stakeholder has a concern regarding a decision not to do a sensitivity analysis in connection with any specific risk assessment, the stakeholder should raise that concern with OPP.

Comment: One commenter believed the Agency should conduct its sensitivity analyses using the full LOD for acute risk assessments–the most health protective approach – and OPP should request additional data or improved analytical methods if the analysis shows the assessment is sensitive to the use of the LOD. "In any case," the commenter stated, " the burden to show, based on sound science and data, which nondetected residue samples should be incorporated into the risk assessment at values less than the LOD, should lie on the chemical manufacturing company seeking a registration or supporting an existing registration."

OPP Response: The comment is somewhat confusing. OPP believes that risk assessments should use the best data available to develop for the risk manager the most

reasonable estimates of exposure and risk, consistent with OPP's goal of avoiding the underestimation of risk. Use of ½ LOD (or ½ LOQ) as part of OPP's "baseline" risk assessment is the procedure which is most consistent with this goal. OPP will continue to conduct its sensitivity analyses for acute assessments, when appropriate, using the full LOD. If OPP finds that there is no significant difference in the high-end exposure and risk estimates when ½ LOD is assumed for treated non-detects vs. when the full LOD is assumed for treated non-detects, OPP will have established that the risk and exposure estimate developed is not sensitive to an assumption concerning the residue concentration in treated non-detects. In this way, OPP will have established that its risk assessment (and any ensuing risk management decision) is *inherently robust* and would not be affected by any reasonable series of true values for the residue concentration which lie below the LOD or LOQ.

Comment: One commenter indicated that while a sensitivity analysis may be a useful tool to determine how sensitive a risk assessment is to the assignment of imputed values to ND samples, the science policy should point out that a sensitivity analysis of the effect of using ½ LOD instead of the full LOD is unlikely to be useful The commenter stated that such a substitution could change the result by 2x at most, and the factor would only be that high if all values were ND. The commenter emphasized that the difference between using ½ LOD and using true zero, however, could be quite significant in some situations.

OPP Response: OPP agrees with the commenter, and now discusses this point in the revised policy document. The following test will be added as a criterion for determining whether a sensitivity analysis is performed--when less than 50% of the aPAD is occupied for all population subgroups of concern, it is unnecessary to perform a sensitivity analysis to verify that the ½ LOD assumption is not potentially underestimating exposure.

D. Guidance on Operational Determination of LOD and LOQ

Comment: One commenter stated that the methods for determining LOD and LOQ are adequately defined for use in this policy, but believed that a specific policy focusing on the methods for determining LOD and LOQ is necessary. The commenter stated that it supported and had used the Keith paper in determinations of the LOD and LOQ.

Several commenters encouraged OPP to seek input from interested parties on the guidance under preparation on the experimental and statistical determination of LODs and LOQs. One commenter emphasized that any new guidance must be reasonable and practical for all residue analysts: industrial, academic, government, etc., and that the innate variability associated with every residue value used in assessing risk should be taken into consideration and guidance for stringent statistical determinations of LOQ and LOD should be based on the practical application of any tolerance enforcement method.

OPP Response: To date, OPP has not issued formal guidance or suggested/recommended procedures, or made available a list of acceptable methodologies for the estimation of LOD and/or

LOQ values for pesticide residue analyses. Due in part to the many valid operational definitions of LOD and LOQ and procedures used to estimate these, OPP believes it unwise to prescribe any one specific procedure or protocol as a standard universal requirement for pesticide registration submissions. Any reasonable, generally recognized statistical procedure may be considered and will be evaluated. OPP recommends that registrants and other data submitters fully document the procedures and protocols used to estimate the LOD and/or LOQ for review by OPP. Nevertheless, in the interest of informing registrants and other data submitters of at least one method for LOD/LOQ determination which has been acceptable in the past, an Appendix to this Science Policy document has been added. The Appendix is a slightly modified form of a procedure used by USDA's IR-4 program and is published, in part, in 40 CFR (as 40 CFR Part 136, Appendix B).

E. Additional Refinements, Suggestions, and Clarifications

Comment: One commenter stated that Sections II.C and III.C of the document appear to distinguish between data sets containing less than 10-15% non-detects and data sets containing between 15-50% nondetects. In the former case, the commenter stated that the ½ LOD (or ½ LOQ) substitution method appears to be acceptable to OPP while the latter case (with 15-50% nondetects) seems to be most appropriate for statistical imputation procedures. If this is indeed the case, the commenter stated that the proposal appears really to be intended for data sets that contain more than 50% nondetects and that this should be clarified.

OPP Response: In general, OPP considers that the "replacement" or "substitution" method (replacing treated non-detects with ½ LOD or ½ LOQ) will result in reasonable estimates of risk and exposure if the number of non-detects is small (e.g, 10-15%). Registrant's are encouraged to use the substitution method in these instances and OPP would perform sensitivity analyses in these situations only on a case-by-case basis. When the number of non-detects increases to greater than ca. 10-15% (but is still less than 50%) risk assessments should be performed using the replacement method, and the effect of the substituted values should be assessed by performing a sensitivity analysis, to determine whether the relevant risk and exposure estimates are significantly affected. Such an analysis should be included as part of the risk characterization. If it is determined that the effect of this substitution is significant, it may be desirable to use statistical methods developed for censored data (as explained in Part III of the revised document). When data sets consist of >50% non-detects, the handling of ND's should be considered on a case-by-case basis and no general policy exists. As summarized in EPA's *Guidance of Data Quality Assessment: Practical Methods for Data Analysis* (EPA/600/R-96-084) (US EPA 1998):

> All of the suggested procedures for analyzing data with nondetects depend on the amount of data below the detection limit. For relatively small amounts below detection limit values, replacing the nondetects with a small number and proceeding with the usual analysis may be

satisfactory. For moderate amounts of data below the detection limit, a more detailed adjustment is appropriate. The situations where relatively large amounts of data below the detection limits exist, one may need to consider whether the chemical was detected as above some level or not. The interpretation of small, moderate, and large amounts of data below the DL *[detection limit]* is subjective. Table 4.7.1 provides percentages to assist the user in evaluating their particular situation. However, it should be recognized that these percentages are not hard and fast rules, but should be based on judgment.

In addition to the percentages of samples below the detection limit, sample size influences which procedures should be used to evaluate the data. For example, the case where 1 sample out of 4 is not detected should be treated differently from the case where 25 samples out of 100 are not detected. Thus, this guidance suggests that the data analyst consult a statistician for the most appropriate way to evaluate data containing values below the detection limit.

Comment: One commenter indicated that the considerations in Section II.G. appear to apply equally to all analytes of concern, whether they are metabolites or parent molecules and suggested revising the header of that section to substitute the term "Analytes" for "Metabolites."

OPP Response: OPP agrees and the revised document has been changed to reflect this suggestion.

Comment: One commenter stated that EPA should consider using residue disappearance kinetics from such crops to predict concentrations which would be expected at harvest rather than utilizing half of the LOD for such chemicals. The commenter acknowledges that such kinetics would have to be established at measurable concentrations of the residues.

OPP Response: OPP has issued guidance concerning predicting residue concentrations from residue disappearance data. This guidance is currently available on the web (see "Guidelines for the Conduct of Residue Decline Studies for Use in Probabilistic Risk Assessment" at http://www.epa.gov/oppfead1/trac/science/#additional) and is expected to be announced to be available as a revised policy in the Federal Register in March, 2000.

Comment: One commenter stated that the science policy document would be strengthened by adding specific examples of how certain aspects of the policy would work in practice.

OPP Response: OPP had moved an illustrative example of calculation of a weighted ½ LOD from the document dealing with censored data to a more appropriate location in the combined, revised document. This was done in response to a specific suggestion from a commenter and is believed to strengthen the document as a whole by illustrating how this aspect of the policy would work in practice. OPP has also added an example calculation of Cohen's procedure to the section on censored data methods.

Comment: One commenter cautioned the Agency that when relying on percent crop

treated data for assigning "true zeroes" to a portion of the residue samples below the LOD, the Agency must be diligent about verifying annually that the percent crop treated is consistent with the risk assessment and, if not, modify the risk assessment. The commenter stated that it is unclear, to date, how and whether the Agency will accomplish this task. The Agency guidance, the commenter continued, indicates that the percent crop treated will be decided on a "case-by-case" basis considering such factors as whether the percent of crop treated fluctuates from year to year or whether there is an increasing or decreasing trend. However, the commenter noted, the Agency fails to indicate what level of stability in the percent crop treated is sufficient for using a fixed percentage in calculating the number of non-detected samples that are true zeros.

OPP Response: The Agency is cognizant of the requirement that the risk assessment be consistent with the percent of crop treated. OPP currently uses a high-end percent crop treated estimate from its Biological and Economic Analysis Division which is unlikely to underestimate the true national percent crop treated and believes that this is also unlikely to underestimate the percent crop treated in localized geographic areas.

Comment: One commenter pointed out that the varying sensitivity of analytical techniques used by residue laboratories will impact the "weighted average" value applied by EPA to nondetections associated with a treated crop. In addition, substantial variation in the LOD's for a single crop/chemical combination will preclude the use of statistical estimation (Cohen's Method) of the adjusted mean residue value. The commenter, as a result, encouraged OPP to work with registrant's laboratories, state regulatory agencies, and other federal agencies that monitor pesticide residues to establish consistent and analytically sensitive methods for detecting residues to reduce the level of variation in LOD's and therefore reduce the uncertainty in EPA's pesticide risk assessments.

OPP Response: OPP is working with USDA's PDP program in an attempt to obtain more consistent LOD's across state and federal laboratories which contribute to the PDP program and encourages registrants and other data submitters to produce data which meet or exceed the quality standards routinely seen in PDP data. In general, OPP finds the LODs which the USDA's PDP program can routinely obtain to be entirely adequate to serve as a basis of OPP risk assessments.

III. Response to Comments on Censored Data Issues

A. Restrictions and Requirements for Use of Methods

Comment: One commenter supported the use of statistical methods to estimate mean residue values accounting for data points below the LOD for use in chronic risk assessments when this is statistically appropriate. The commenter stated that in Monte Carlo assessments, estimated individual data values should be used, not mean levels, but that if those values can be extrapolated based on a known and demonstrated residue distribution, this statistical technique is acceptable

and appropriate for acute risk assessment. The commenter was concerned, however, that EPA has not emphasized enough in its guidance document that the underlying residue distribution must be *demonstrated* to be normal or lognormal in order to use Cohen's method. The commenter stated that while EPA mentions that normality tests must be conducted, it later refers to the "assumed normal distribution." The commenter believed that it is incumbent upon the registrants to provide evidence that the residue data are normally distributed or can be transformed into a normal distribution and that such evidence must be based on actual data, not on assumed or surrogate residue distributions. Furthermore, the commenter stated that all data points must be used to demonstrate that the residues are normally distributed and that outliers should not be omitted simply because they do not follow the expected distribution.

OPP Response: OPP is in agreement with the commenter. The reference to an "assumed normal distribution" was made to reflect a common statistical convention that one cannot *prove* a given distribution belongs to a hypothesized family of distributions (e.g, normal, lognormal, Poisson, etc.), but rather can only provide sufficient evidence to suggest that the hypothesized distribution is not inconsistent with actual distribution (analogous to either "rejecting or "failing to reject" a hypothesis). If there is insufficient evidence to demonstrate that a distribution is not normal, then it is reasonable to refer to it as an "assumed normal" distribution.

Comment: One commenter questioned the need for at least 50% of the data to be detects in order for "left-data uncensoring" to be conducted. He indicated that this, at least in part, appeared to be linked to EPA's concern that the data be shown to be normally or lognormally distributed. The commenter believed that this requirement needed to be more fully detailed from a science standpoint. The EPA policy paper itself, the commenter remarked, cites Ott (1995) in stating that a lognormal distribution of pesticide concentrations in food crops can be supported on a theoretical basis. Indeed, the commenter state, findings of U.S. EPA's Office of Research and Development support the use of lognormal distributions for environmental concentrations (Singh et al., 1997). Selier and Alvariz (1996), he continued, present arguments showing that if data are limited, the selection of a lognormal distribution is a sound, rationale choice for environmental contaminants. In light of these observations, the commenter claimed that a requirement demonstrating that a distribution is normal or lognormal is unnecessary. Since a lognormal distribution of data can be reasonably assumed and supported, the commenter continued, the concern of the policy paper might be whether there is sufficient data richness to understand the degree to which the data are skewed. If this is the case, the commenter stated, it should be clearly presented as such as should the methodology for predicting and dealing with left-skewed data. The commenter also stated that the possibility of tying the allowable proportion of non-detects to the level of confidence in the assumed shape of the distribution should also be considered.

OPP Response: OPP believes that the assumption of a lognormal distribution for environmental concentrations is not unreasonable, but will continue to recommend that the registrant demonstrate that a lognormal distribution provides a statistically adequate fit to the data (or, more accurately, that a lognormal distribution is not inconsistent with the data). The commenter is correct in that one concern of the policy paper is whether there is sufficient data

richness to understand the degree to which the data are skewed. OPP believes that this will necessarily be determined on a case-by-case basis taking into consideration a variety of factors which impinge on the data including sampling effort, degree of skewedness, percentage of nondetects, and limits of detection/quantitation. No general policy guidance can be established.

Comment: One commenter expressed disappointment concerning the requirement that a sufficient portion of the data set be at measurable levels in order for "data uncensoring" to be used. He believed that this is a severe limitation to the broad-based utilization of the proposed methodology for pesticide food residue exposure assessment since for many residue data sets nondetects comprise a substantial portion of the data. The commenter stated that since it is possible to statistically evaluate the goodness of fit of a distribution, statistical criteria rather than the percent NDs should be the established approach for use of an uncensoring method.

OPP Response: In the revised version of its science policy paper, OPP has made clear that this document makes recommendations. It does not establish requirements with respect to the percentage or absolute number of detects. Moreover, the document makes clear that for those instances where more than 50% of the data are non-detects and/or less than 10-20 data points are available, that this will be judged on a case-by-case basis. OPP does not intend this to be a limitation to the broad-based utilization of the proposed methodology but will, however, insist that any utilization of such data imputation procedures be adequately supported from both a statistical and data quality perspective.

With respect to the commenter's suggestion that it is possible to statistically evaluate the goodness of fit of a distribution (implying that it is not necessary to be concerned about how much data are available), OPP reiterates the following cautions with regard to statistical goodness-of-fit tests which are described in OPP's Monte-Carlo Guidance document (US EPA 1998b):

> Goodness-of-fit tests are formal statistical tests of the hypothesis that the set of sampled observations are an independent sample from the known or assumed distribution. The null hypothesis, *H* , is that the randomly sampled set of observations are independent, identically distributed random samples from a population with the hypothesized distribution. The GoF tests indicate whether the hypothesized distribution can be reasonably rejected as improbable. It is important to recognize that failure to reject H is not the same as accepting H as true. These tests, taken alone, are not very powerful for small to moderate sample sizes (i.e., subtle but systematic disagreements between the data and the hypothesized distribution may not be detected); conversely, the tests can be too sensitive for large numbers of data points -- that is, for data sets with a large number of points, H will almost always be rejected. Care must be taken not to over-interpret or over-rely on the findings of goodness-of-fit tests. It is far too tempting to use the power and speed of computers to run goodness-of-fit tests against a generous list of candidate distributions, pick the distribution with the "best" goodness-of-fit statistic, and claim that the distribution that fit "best" was not rejected at some specific level of significance. This practice is statistically incorrect and should be avoided [Bratley *et al.*, 1987, p 134]. ...Goodness-of-fit tests have notoriously low power and are generally best for rejecting poor distribution fits rather than for identifying good fits. For small to medium sample sizes, goodness-of-fit tests are not

> very sensitive to small (but potentially significant) differences between the observed and fitted distributions. On the other hand, for large data sets, even minute differences between the observed and fitted distributions may lead to rejection of the null hypothesis. For small to medium sample sizes, goodness-of-fit tests should best be viewed as a systematic approach to detecting gross differences.
>
> We note that there is absolutely no substitution for careful visual inspection of both the data and the theoretical distribution of the fit to the data. The human eye and brain are able to interpret and understand data anomalies far beyond the ability of any computer program or GoF tests. GOF tests may, *at best*, simply serve to confirm what the analyst has found though visual inspection. One may quite appropriately decide to retain a particular probability model despite having rejected it on the basis of GoF tests if it appears to be a good fit to the data as judged by the visual inspection of the probability plots and other comparisons.

Comment: One commenter indicated that the policy issue is confusing as to the Agency's intent regarding the nature of data sources that can be used and how they can be used in statistical treatment of left-censored data and there is a need for clarification in a manner that is unambiguous and clearly justified. The commenter mentioned that residue data for blended commodities, data bases with several different LODs, and monitoring data are all described in contexts that indicate a restricted willingness on the part of OPP to employ statistical tools.

OPP Response: OPP has attempted to provide clarification of the issues in its revision.

Comment : One commenter indicated that it is not necessary that Cohen's method be restricted to cases where no more than 50% of the data are below LOD, stating that the performance of the method depends on the absolute number of observations above the LOD in addition to the percentage of observations above the LOD. Thus, the commenter believed that the adequacy of the method should be decided on a case-by-case basis.

OPP Response: OPP has revised the document to state that statistical treatment will not be necessarily discouraged in those situations in which more than 50% of the data are (treated) non-detects. As stated earlier, the revised version contains recommendations with respect to the percentage or absolute number of detects. It states that, in those instances where more than 50% of the data are non-detects and/or less than 10-20 data points are available, the appropriateness of statistical treatment will be judged on a case-by-case basis. However, OPP reiterates that any utilization of such severely censored data must be adequately supported from both a statistical and data quality perspective and extensive characterization and sensitivity analyses will likely be necessary.

B. Other Methods

Comment: One commenter believed that the science policy document, as presented, severely restricts the available methodologies and the conditions under which imputing may be usefully applied. The commenter states that several long-standing methods are available that are

less restrictive in their applicability than are Cohen's method. The commenter stated that although the policy paper points in the proper direction for use of statistical techniques in the conduct of more scientifically sound risk assessments, the science will be little advanced if implemented as stated. For example, the commenter stated that sources cited in the policy paper draw attention to numerous methodologies that have been successfully employed to determine the likely distribution of environmental residues reported as ND. The source guidance on which this policy paper is based (*Guidance for Data Quality Assessment: Practical Methods for Data Analysis*: US EPA 1998a) suggests several methods in addition to Cohen's method including the methods of trimmed mean and Winsorized mean and SD (both of which can be used if the percentage of ND's are between 15-50%) and a Test for Proportions (which can be used if the percentage of ND's are between 50-90%). The commenter stated that it was noteworthy that this methodology (1) provides methods that can be used across a wide range of ND's in the data base used; (2) allows for several alternative options of statistical method within a given range of NDs in the data base used; (3) describes ranges of NDs for use of the Cohen Method and the use of ½ LOD/LOQ default values different from the science policy paper and its companion paper on default values.

Along similar lines, another commenter indicated that additional methodologies should be sought or included:

> The choice of a given method may not only relate to the population of ND's in the data set but the way transformed data are to be used (to better predict the mean in an analysis utilizing AR values or the distribution of ND's in a Monte Carlo analysis). To restrict OPP methodology for human dietary exposure assessment to a single technique and somewhat rigid conditions...seems inappropriate and needs to be clarified by EPA, especially in light of the seeming inconstency in use of methodology as presented in the science policy document and the source guidance document.
>
> Broad scope assessment of various imputing techniques for dealing with NDs in environmental sampling have been reported in the literature. A review of methods for imputing left-censored data shows a variety of techniques based on maximum likelihood analysis (MLE), iterative techniques, regression on order (ROE), bias-corrected MLE, restricted MLE, modified probability plotting, Winsorization, and lognormalized statistics (EPA D-LOG). Berthouex (1991) found Cohen's MLE and ROE methods equally useful. Modified probability plotting (Helsel's Robust Method) was developed from Helsel's 1990 review and is a favored default method in the program UNCENSOR (Newman et al., 1995). We have performed our own in-house evaluation of these various statistical tools using both UNCENSOR ver 4.0 as well as Mini-tab based tools. In our view, Helsel's Robust Method lives up to its name.
>
> In short, while we and others find little argument in the use of Cohen's MLE method, the acceptable conditions of use as outlined by the Agency appear very limited. Numerous alternative techniques can lead to more robust solutions that can be more flexibly applied to a broader variety of conditions than the very limited ones suggested by the science policy paper. The fact that statistical tools are available to evaluate the goodness of fit of modeled to observed distributions affords utility in the wider range of uncensoring methods under a broader range of conditions than afforded in the policy. USEPA OPP should further evaluate approaches and

provide the scientific rationale underpinning the methodology and approaches that have been put forward.

OPP Response: The policy document has been revised to clarify that the discussion of methodologies is designed to be illustrative and does not limit what methodologies can be used. OPP is not attempting to preclude the use of any given method, but rather attempting to establish guidance for "good practices" with respect to use of these methods. Additionally, the discussion of methodologies has been considerably expanded to include mention of several other potentially useful methodologies.

Comment: One commenter suggested that an improved alternative to Cohen's method is the Robust Method as described by Helsel (1990). In this method, the distribution below the LOD is estimated by using replacement values from a fitted distribution below the LOD. The estimated mean residue is calculated using actual observations above the LOD (not from a fitted distribution) and the replacement values below the LOD. The commenter claimed that Helsel's method is superior in that values above the LOD are not replaced by the fitted distribution whereas in Cohen's method only a calculation of the mean of the fitted distribution is provided. The commenter continued by stating that the robust method can be used with non-normal distributions and with data having more that one LOD. The fitting of the distribution can be performed using maximum likelihood estimation technique. ccc

OPP Response: OPP has investigated Helsel's method and believes in many cases that this is an appropriate method to use. The document has been revised to reflect this.

Comment: One commenter stated, "Cohen's method is a long standing and clearly recognized technique for treatment of left-censored data"... but "is three decades old and considerable effort... has been extended in developing alternative or modified methods that are more robust than Cohen's method for dealing with environmental data." The commenter indicated that OPP seems to have concerns about the types of data sets for which this method can be reliably used, but has not presented their rationale for restricting statistical treatment of left-censored data to Cohen's method conducted under tightly prescribed conditions. The commenter stated that "in light of the substantial body of literature pointing to alternative methods that are robust to conditions less restrictive than Cohen's MLE procedure, the EPA OPP should develop a stronger scientific rationale for the use of this particular method with the conditions proscribed in the science policy paper vs. alternatives used under a variety of use conditions."

OPP Response: In the revised version of the document, OPP does not recommend that the statistical analysis of left-censored data using Cohen's method be limited to certain prescribed conditions. Nevertheless, any method which a registrant or other data submitter uses to impute values for below detection limit data must be adequately supported by both a sufficiently rich data set above the detection limit and a statistically-robust methodology for imputing those values. It should be remembered that, under many of the conditions which statistical estimation procedures are used to impute these values, a previous sensitivity analysis would have been performed which would have demonstrated that the relevant risk and exposure estimates are dependent upon (and

significantly affected by) the assumptions made for the concentrations represented by the <LOD or <LOQ values. That is, a sensitivity analysis would have previously been performed, the result of which would have shown that the use of 0 vs. ½ LOD or the full LOD vs. ½ LOD has a significant effect on the risk or exposure estimate.

Comment: One commenter stated that he believed that OPP's proposed approach is "well short of state of the art." Instead, the commenter suggested the following general strategy:

- If less than 10% data are not quantified, use half the LOD or LOQ as appropriate and do statistics using the normal formulae
- If more than 10% of the data are censored, but there is only one censoring value, use regression on normal scores
- For complex situations where there are multiple censoring values, use the method of maximum likelihood.

One commenter believed that OPP should permit the use of statistical methods other than Cohen's method. The commenter provided seven other potential methods which are described in Newman et al. and are available in shareware UNCENSOR 4.0.

OPP Response: OPP has considerably expanded the number of statistical methods recommended for consideration including regression on normal scores and use of maximum likelihood methods. Whichever statistical methods are selected for use, they must be adequately supported and justified.

Comment. One commenter recommended that if Cohen's method or any other method is used to fit a distribution to the observed data, then the estimate of the mean should be computed directly from the observed values above their detection limits and replacement values for the values below their LOD's.

OPP Response: Although the commenter states that this situation may occur, no reference or documentation was cited. In any case, if the predicted distribution above the LOD/LOQ (using, e.g., Cohen's method) differed significantly from the observed distribution above the LOD/LOQ, OPP would further investigate the assumptions implicit in the data uncensoring method before proceeding further with its analysis.

C. Suggested Clarifications/Corrections

Comment: One commenter indicated that while Cohen's method readily provides a mean and standard deviation which are appropriate for assessing chronic pesticide food residue risk, it would be helpful to provide additional guidance on how Cohen's method may be used to impute discrete values for use in acute pesticide food residue risk.

OPP response: The commenter is correct. Details of this how this could be performed now have been included in the revised version of the document. Briefly, Cohen's method could be used to determine the mean and standard deviation of a distribution as described in the document. This now defined distribution could then be used to impute those values which lie below the detection limit by using the inverse cumulative distribution function, Φ^{-1} , where $\Phi^{-1}(p)$ = z-score and p = *n*/N+1 (assuming a normal distribution). Any imputed non-detect values used to "fill-in" the distribution (i.e., replace the <LOD values with more appropriate single-valued finite estimates) would be calculated as follows:

"Fill-in" value = z-score * SD + mean.

For example, suppose that there are 100 data points of which 98 are above the detection limit and two are below the detection limit. Further supposed that calculation via Cohen's method results in an estimated mean of the distribution of 10.0 and a standard deviation of 2.0. The values would then be ranked, and (since there are 100 total values), the first non-detect would occupy the 0.01 quantile and the second non-detect would occupy the 0.02 quantile. The corresponding p values would be calculated as p_1= 1/(100+1) = 0.0099 and p_2=2/(100+1) = 0.0198 for the first and second ND values, respectively. Using a normal probability table, one would determine that $\Phi^{-1}(p_1) = \Phi^{-1}(0.0099) =$ -2.33 and $\Phi^{-1}(p_2) = \Phi^{-1}(0.0198) =$ -2.06. The fill-in values associated with these two z scores are -2.33*(2.0) + 10.0 = 5.34 and -2.06*(2.0) + 10.0 = 5.88. Thus, 5.34 and 5.88 would be the two fill-in values associated with the two non-detects.

Comment: One commenter expressed concern about the lack of specificity EPA provides in this guidance regarding sample size requirements for estimating an adjusted mean or imputing residue values below the LOD. EPA provides little information, the commenter continued, on the importance of sample sizes in estimating values for non-detectable residues. While the commenter claimed to support EPA's proposal to allow for extrapolation of individual values below the LOD for use in acute risk assessments, the commenter recommended that EPA provide more complete guidance on this method to better inform the public.

OPP Response: OPP has inserted the following information in its revised document:

> Additional recommended criteria for use of Cohen's methodology is that not more than 50% of the data set be censored (ideally, less than 20% should be censored) and/or at least 10 non-censored data points (with 20 or more being strongly desirable) be available. Exceptions to these recommended criteria can be made on a case-by-case basis. However, with respect to the exceptions, it should be remembered that in many cases it is likely that a more refined estimation procedure such as Cohen's method is being using precisely because the insertion of ½ LOD for ND residues resulting in unacceptable risk while the substitution of 0 ppm for ND's resulted in risks below OPP's level of concern. That is, in many cases Cohen's method will be used because OPP's risk estimate or resulting decision is very sensitive to assumptions about values to assign to ND residues. Thus, OPP feels justified in recommending stricter criteria for use of Cohen's method than might normally be used in attempting to estimate a "best" estimate of a

mean residue value.

As emphasized in the above inserted paragraph, this is guidance only and exceptions will be considered on a case-by-case basis.

Comment: One commenter recommended that the paper more clearly identify the techniques referred to as *imputing of left-censored data* in that it describes statistical methodology for dealing with ND's using some form of data extrapolation. As stated by the commenter:

> Within dietary risk assessment, there is interest, additionally, in statistically imputing single-serving residue values from averages of mixed commodity data; this is more of a data deconvolution exercise. Similarly, construction of empirical probability functions (EDF) is a form of imputation as well that results in data interpolation. Since each technique at one time or another is discussed within the context of imputing data for dietary risk assessment, it is best in the present paper to clearly identify imputing (or "uncensoring") of left-censored data.

OPP Response: OPP agrees with the commenter and has made the requested clarifications.

Comment: One commenter pointed out several errors in the mathematical equations set forth in the Science Policy in which a square value was mistranscribed and in which a parentheses was misplaced.

OPP Response: These errors have been corrected in the revised document.

IV. List of Commenters

Byrd, Daniel. 1999. CTRAPS (Consultants in Toxicology, Risk Assessment, and Safety). February 4.

Craigmill, Authur L. and Machael A. Kamrin, EXTOTOXNET and Michigan State University. 1999. February 2.

Franklin, C.A. 1999. Health Canada. February 1.

Friedman, Priscilla L. 1999. DuPont Agricultural Products. February 3.

Gersich, F. M. 1999. Dow Agrosciences. February 3.

Gustafson Chemical Company. 1999. February 4.

Ginevan, Michael E. 1999. M.E. Ginevan and Associates. February 16.

Henry A. Wallace Institute for Alternative Agriculture. 1999. February 3.

Kenney, Jeannine. 1999. Consumers Union. February 4.

Maslyn, Mark A. 1999. Implementation Working Group. February 4.

Moore, Sam. 1999. Kentucky Farm Bureau. February 5.

Novigen Sciences, 1999. February 18.

Whitacre, David. 1999. Novartis Crop Protection. February 4.

V. Bibliography

Helsel, DR. 1990. Less than obvious: statistical treatment of data below the detection limit. *Environ. Sci Technol.* 24(12): 1766-1774.

Ott, Wayne R. 1995. Environmental Statistics and Data Analysis. Lewis Publishers (New York: New York).

U.S. EPA. 1998a. Guidance for Data Quality Assessment: Practical Methods for Data Analysis EPA QA/G9. QA-97 Version. Office of Research and Development. EPA/600/R-96/084. (available in downloadable Adobe Acrobat pdf format on the world wide web at http://www.epa.gov/r10earth/offices/oea/epaqag9.pdf)

U.S. EPA. 1998b. Draft document. Guidance for Submission of Probabilistic Human Health Exposure Assessments to the Office of Pesticide Programs. November 4. (available in downloadable Adobe Acrobat pdf format on the world wide web at http://www.epa.gov/fedrgstr/EPA-PEST/1998/November/Day-05/o-p29665.htm)

U.S. EPA. 1999. Office of Pesticide Programs. Health Effects Division. "Classification of Food Forms With Respect to Level of Blending." HED Standard Operating Procedure 99.6 (8/20/99) August 20.