APPENDIX I

A Review and Critique of the Quantitative Risk Assessment for Polychlorinated Biphenyls Proposed by the Health and Environmental Review Division (HERD) of the U.S. Environmental Protection Agency

by

Dr. Raymond D. Harbison and Dr. Robert C. James

Division of Interdisciplinary Toxicology
University of Arkansas for Medical Sciences
Little Rock, Arkansas 72205

January 1984

Re: Document No. DPTS-62032
Submitted by Monsanto Company
February 3, 1984

The following report reviews and critiques three separate documents developed or used by the Health and Environmental Review Division (HERD) from the Office of Pesticides and Toxic Substances of the U.S. Environmental Protection Agency. The documents HERD relied upon or developed for their assessment of the risks associated with PCB exposure were:

1. Carcinogenic Risk Assessments of Polychlorinated Biphenyls (PCBs), Health and Environmental Review Division, Office of Toxic Substance, September 1 (1983), Office of Pesticides and Toxic Substances, U.S. Environmental Protection Agency, Washington, D.C. 20460.
2. Quantitative Risk Assessment of Reproductive Risks Associated with Polychlorinated Biphenyl (PCB) Exposure, Health and Environmental Review Division, Office of Toxic Substances, September 1 (1983), Office of Pesticides and Toxic Substances, U.S. Environmental Protection Agency, Washington, D.C. 20460.
3. Exposure Assessment for Incidentally Produced Polychlorinated Biphenyls (PCBs), Draft Final report of Appendix A: Generic Exposure Scenarios, EPA Contract No. 68-01-6271, Task 21, Versar Inc., 6850 Versar Center, Springfield, VA, August 15 (1983).

A Review and Critique of the Quantitative Risk Assessment for Polychlorinated Biphenyls Proposed by the Health and Environmental Review Division (HERD) of the U.S. Environmental Protection Agency

This review of the September 1, 1983, risk assessment documents prepared by the Health and Environmental Review Division (HERD) of the Office of Pesticides and Toxic substances on polychlorinated biphenyls (PCBs) summarizes many concerns as succintly as possible. The first section of the review contains only a few general comments and a brief discussion of our position on the HERD risk assessments. The second section contains specific comments on three documents labeled as Appendices A, B and C. In the second section we have attempted to point out some but not all of the inconsistencies, contradictions and problems associated with the HERD cancer risk assessment, the HERD reproductive risk assessment, and the Versar Corporation's exposure assessment of incidentally produced PCBs. The reviewers recognize the difficulties associated with any risk assessment, yet were particularly disturbed and distressed with the failure of HERD to rely on factual information and generally accepted scientific principles and concepts. With this in mind it is hoped that the criticisms raised concerning the HERD risk assessments will be utilized by HERD in a constructive manner to revise their approach.

General Comments:

1. In their present form both documents are poorly written and contain numerous typographical errors. It appears that the documents were not

reviewed as many statements are confusing, ill-conceived and poorly substantiated. Because of the poor quality of the present documents it does not seem conceivable that they could be used for any purpose in their present form. Therefore, it is strongly recommended that both documents undergo internal editorial examination and revision before they are again released.

2. Both documents make use of numerous unsupported or poorly substantiated claims. Some of the statements seemed to be nothing more than a rationalization of the weaknesses of the approach taken by HERD while others tend to be misleading as they represent the feelings of the HERD group rather than those of the scientific community that they should reflect. These severely undermine the credibility of the risk assessments and the HERD group itself. For example, in the carcinogenic risk assessment document HERD proposes that one important reason for using the NCI data over the Kimbrough data is that its lowest dose is closer to the exposure estimates thereby instilling greater confidence in the extrapolations. Yet in actuality there is essentially no important difference between the 1.25 mg/kg/day dose used in the NCI study versus the 4.9 mg/kg/day dose used by the Kimbrough study when one compares these doses to the exposures estimates many of which are more than 1,000,000 times lower. In another place HERD suggests that since the risk estimates of these two cancer bioassays showed some agreement that this was evidence the various strains reacted similarly to the PCB carcinogenic insult. Since the Kimbrough study was positive and the NCI study negative we find such an insinuation to be blatantly false. The similarities occur, as HERD fails to point out,

only because of the manner in which HERD has mathematically manipulated and modeled the negative NCI study.

3. In several places the documents clearly illustrate the inexperience and lack of knowledge of the persons involved with the risk assessments. The persons performing the HERD risk assessments do not understand the biologic responses they are modeling nor the limitations of some of the models they are using.

4. These two risk assessments combine an additional uncertainty factor that many risk assessments do not. Rather than estimating safe exposure limits or estimating risks for actual exposures, these documents used estimated exposures generated by a separate document to then estimate the risk. Thus, the risks generated suffer from inaccuracies and uncertainties inherent to both the risk and exposure modeling. Some of the assumptions made in the Versar document to estimate exposures are scientifically unsupportable while others are devoid of reality or credibility. Uncertainties in the exposure estimates compound the the uncertainty associated with the risk modeling and further undermine the credibility and utility of these risk assessments.

5. The purpose of any risk assessment is to provide a basis for or a guidance in decisions concerning the use of or exposure to chemicals. The HERD document should not be used as a basis for regulatory decisions or as guidance in decisions. So many unrealistic assumptions have been made in both the exposure and risk estimation models that these risk assessments provide, at best, only a weak relative rank ordering for the various "proposed" exposure scenarios. As far as being used to determine whether certain exposures have any health

impact, there are far too many uncertainties in the calculations for them to be seriously relied upon. For this reason an obvious question should be asked, "Why were these risk assessments developed?". In 1981 the Office of Technology Assessment published a fairly thorough appraisal of the assessment technologies used for estimating chemical-induced cancer risks (see Assessment of Technologies For Determining Cancer Risks From the Environment, OTA, 1981). In this publication and from many scientific publications on the subject, it is clear that risk estimates are model dependent and can vary by several orders of magnitude depending on which model is used. HERD aptly re-demonstrated this fact in both documents. There is also no way of knowing which estimate, by which model is the most accurate reflection of risk. So with no new data, and no better foundation for using any one particular model, HERD has re-assessed the risks associated with exposure to PCBs without being able to provide any credible reasons that their estimates are any more valid than those generated in previous risk assessments. That this would be true regardless of how their assessment might have been conducted raises questions regarding HERD's intentions and their understanding of risk assessments in general. In addition to this, HERD's risk assessments are so badly flawed that concern is further increased for HERD's wasted efforts on this particular subject.

6. A fact that probably best highlights the weakness of HERD's approach for these two risk assessments is the fact that HERD's risk estimates for reproductive problems is slightly greater at all exposure estimates than their cancer risk estimates. This is largely due to the

fact that they used a more conservative mathematical model to estimate reproductive risk than they used for carcinogenic risk. Such an approach cannot be defended by HERD and will not be accepted by the scientific community. Furthermore, HERD's approach to modeling reproductive risks is one in which a dose 100,000 times lower than one actually producing an undesirable effect (in this case infant mortality) is expected to produce the same effect once in every 100,000 exposures. This is not an approach that will receive much, if any, support from the scientific community and as such only further undermines any confidence others might have in HERD's efforts.

Since this review of the HERD risk assessments has found them to be of an unsatisfactory quality and nature, it seems appropriate to reiterate some recent and specific points made by the National Research Council of the National Academy of Sciences concerning risk assessments and the regulatory process (Risk Assessment In the Federal Government: Managing the Process, National Research Council, National Academy Press, 1983). In this book, on page 131, are listed the problems that have been associated with federal regulatory agency risk assessments. These are:

- Bias - i.e. regulators may skew their risk assessments of a particular substance to support a regulatory preference.
- Exaggeration - i.e. that the instinct to support a position with every available argument may distort interpretation of scientific data, choice of extrapolation procedures, and assumptions about human exposure.
- Poor Public Understanding - i.e. if risks are misdescribed it follows that public perception of the risks will be inaccurate.

- Poor Quality Personnel - i.e. that regulatory agencies cannot attract or retain adequate numbers of highly qualified scientists to perform risk assessments.
- Inconsistency - i.e. that agencies have applied inconsistent criteria and reached inconsistent results in assessing the risks posed by the same hazards.
- Redundancy - i.e. that different regulatory agencies are concerned with the same hazards thus assessments should be centralized.

Most of the above six problems caused by using diversified agency risk assessments in general can be specifically applied to the HERD PCB risk assessments as well. Below the criticisms are again listed with specific examples of each problem that applies to the HERD documents. These are:

- Bias - Both risk assessments are based on faulty assumptions. In the carcinogenic risk assessment HERD utilizes a negative study to develop risk estimates. In the reproductive risk assessment HERD uses a mathematical approach for carcinogens that is more conservative than the model used to estimate carcinogenic risks. Further, HERD argues that safety factor approaches should not be used to calculate reproductive risks then uses an approach which is equivalent to using a variable safety factor on a linear scale. HERD can provide no evidence that reproductive toxicity at low doses can be reduced to a probability type of event with the outcome equivalent to responses observed with higher, more toxic doses. Yet by modeling the risk in this manner, HERD "skews" the results to much lower values than would be obtained otherwise.

- Exaggeration - HERD concludes that the NCI and Kimbrough et al. cancer bioassays gave similar carcinogenic responses. Nothing could be farther from the truth as one was reported as positive for liver tumors while the other was reported as negative for any form of cancer. HERD also distorts what the Gaylor/Kodell linear interpolation model represents as it is a model in which the existence of a threshold is denied while HERD implies otherwise. In neither risk assessment does HERD provide substantial or even convincing arguments that either assessment is better than any performed previously by the EPA. Yet mention is made by HERD that previous risk assessments were poorly justified and documented. Lastly, in several instances it is obvious that the exposure estimates used are unrealistically high.
- Poor Public Understanding - HERD has repeatedly failed to present the public with accurate and complete facts. For example, HERD has failed to recognize certain scientific principles, such as epigenetic mechanisms for carcinogens and that these carcinogens represent qualitatively different risks to humans. While much human data in recent years has failed to produce evidence, other than chloracne, to support the concerns raised in animal toxicity tests, HERD has failed to recognize this fact. Now HERD presents the public with the view that a more conservative approach must be taken with reproductive risks than normally taken by EPA for estimating cancer risks. It seems reasonable to conclude that the public might be led to perceive PCBs as reproductive risks at any dose while evidence is available to the contrary.
- Poor Quality Personnel - several comments made in both documents reveal that the HERD group does not understand the methods they are

using or the biologic processes they are modeling. In addition, HERD probably cannot state that they are better qualified than the EPA's Cancer Assessment Group (CAG) yet HERD goes so far as to state CAG's risk assessment was poorly written and concieved. HERD should list the experience and qualifications of its personnel if it truly feels capable of performing risk assessments by itself. While we remain receptive to such disclosure it seems likely that HERD, at present, is lacking in experience and expertise for developing risk assessments of any regulatory importance.

- Inconsistency - The EPA is now proposing two cancer risk estimates for PCBs. One performed by CAG and one by HERD. If these assessments produce the same risk estimates then for what purpose was HERD's efforts? If they are different which EPA group is correct and why? To reiterate, much discussion on cancer risk assessments makes any arguments concerning which risk assessment is best a moot point. Thus, in either situation HERD's efforts appear to be a waste of resources to re-work an old problem and for which their solution appears to add nothing new. Compounding this concern is the fact that the public may now ask which assessment for PCBs is correct? Should HERD argue that their risk assessment represents the better approach, then it becomes certain that a number of the EPA's previous risk assessments for other chemicals have been performed inadequately.

In summary, this review and critique of HERD's PCB risk assessment documents has identified a number of problems which we feel must be addressed. Similar to the National Academy of Science we recognize the difficulties facing regulators and that risk management (i.e. regulations

etc.) does differ from the actual risk assessments (i.e. the scientific estimation of the actual human risks). However, while the EPA is charged with and responsible for protecting public health there is no need to misuse or misrepresent scientific information and scientific principles to accomplish this goal or to support a particular position taken by the Agency. As an agency of the federal government protecting the public interest the EPA should always attempt to present an impartial and thorough discussion of the hazards of any chemical. To overstate, misinterpret or present only one side of any issue is not consistent with this goal. Further, to do so only generates scientifically indefensible positions that undermine the credibility of the Agency. Unfortunately, the HERD PCB risk assessments present the EPA with just such a problem. The documents are not only poorly written but the bias and exaggeration of the scientific information used by HERD is clearly obvious to any informed reader. In addition, in certain places HERD has taken the liberty to disparage the quality of previous risk assessments as well as the knowledge of scientists that would take exception to the approaches used by HERD. Lastly, the assessments raise the question as to who within the EPA is best capable of performing risk assessments, particularly for PCBs. That one group within the EPA is redoing recent work by another EPA group, and that this has occurred without any significant inprovement in our understanding of the toxicity of PCBs or risk assessment methodologies is distressing. Further it suggests that a lack of unity, purpose and scientific interest, exists within the EPA, and that the opinions expressed by one division may represent those of a select few rather than that of the agency as a whole, or

the scientific community it relies upon for support. It is hoped that the EPA, and in particular HERD, will attempt to rectify this problem.

In conclusion, the HERD PCB risk assessment contains numerous typographical errors, misinterpretations of data, misleading or biased statements, citations of their inexperience, and an inconsistent approach to the modeling of data for risk assessments. Therefore, the numbers generated lack sufficient credibility to be of any particular use. HERD is urged to re-evaluate its intent, approach and presentation of these risk assessments.

APPENDICES

APPENDIX A

HERD's Carcinogenic Risk Assessments

of Polychlorinated Biphenyls

Specific Comments:

1. There are numerous typographical errors which make this document at times difficult to read and in some instances leaves portions of the document either open to interpretation or obscures the point the author is trying to make. This document should be revised so as to present the HERD risk assessment in a clear and detailed manner so that the information used, the data manipulations and modeling, and what is stated regarding other risk assessments is not vague or ambiguous. For example:

   a) p. 3, 1st para. - **"This memorandum summarized and pulls together for the first time the information available from previous quantitative risk assessments of PCBs, and updates and fills in gaps in that area on HERD risk assessment of PCBs."**

Who or what is being updated and what specific gap in what area of information is being filled by this document?

   b) p.14, last sentence - **"It is at least known that CAS used a bioconcentration factor of 46,000."**

However in table 1, a reproduction of water quality criteria values suggests that a bioconcentration factor of 31,200 was used.

   c) p. 20, last sentence of para. at the top of the page - **"This test was primarily conducted to see if there would be problems in using the combined data for any malignancy as a basis for risk extrapolation, and it was determined there would be more from a statistical view point."**

This sentence is meaningless and obviously too confusing in its present form.

d) p.24, 1st sentence of last paragraph - "It can be seen from table 11 that the estimates of virtually safe dose at a risk of 10^{-6} (1 in 1,000,000) from the independent background models very a million fold from 10^{-2} for the probit to 10^{-8} for the gamma-multihit."

In table 11 the values for the gamma-multihit is actually 10^{-12} or 4 orders of magnitude lower than stated in the text, thus the values actually vary by 10 billion orders of magnitude.

e) The numbers of total malignancies in table 9 cannot be derived from the rest of the data in table 9 in its present form. If the data provided is totaled, one gets larger or smaller numbers than printed in table 9. Would HERD please clarify this discrepancy and explain where their numbers come from?

2. The text contains numerous misleading statements and unsupported conclusions.

a) p. 2, 2nd para. - "1) The extrapolation distance between the lowest experimental (1.25 mg/kg/day) in the NCI study and most of the exposure estimates are not unusually large. This gives one somewhat greater confidence in the assessment."

The low dose of the NCI study of 1.25 mg/kg/day is not sufficiently different from the 4.9 mg/kg/day dose of the Kimbrough study used by the EPA in the previous Water Quality Criteria Document (Federal Register of Nov. 28, 1980) to warrant this statement. A perusal of the exposure estimates used (table 17) shows that the distance of extrapolation is about 4 to 9 orders of magnitude. That HERD has to extrapolate over such a large distance between the exposures and test data dwarfs the 4 fold difference in doses for the test data itself.

Thus, to insinuate that this small difference instills greater confidence in HERD's assessment is a ridiculous statement. In addition, it cannot be said, as HERD implies, that using the lowest data point available improves the confidence of any risk assessment. In many instances the greatest variability and uncertainty is in the lowest data point. In this particular instance the data point HERD extrapolates from does not even represent a positive response.

The confidence placed in any risk assessment is based upon the validity and accuracy of the data and upon any assumptions made to derive the risk estimates using this data. Confidence in a risk assessment should never be a function of the size of the animal doses used in relationship to the expected human exposure. Furthermore, the HERD document fails to inform the reader in the executive summary that their risk assessment is using data from a negative cancer bioassay whose doses are close to or the same as the single dose used in the Kimbrough study. That the HERD group uses a negative study to calculate a human cancer risk does not inspire confidence in the accuracy of these numbers.

b) p. 2, 2nd para. - **"2) The dose-response data for total malignancies are also linear, this corresponds well with the "linearized" upper 95% confidence limits from the CAG."**

This statement is meaningless and appears to be an attempt by HERD to mislead the reader into thinking that the HERD's use of the NCI data is valid just because it compares well with CAG's initial approach. HERD used three data points ranging only between 1.25 to 5.0 mg/kg/day and which had only modest differences in total malignancies, that HERD then observed that these points could be linearized is not surprising. It can be done for any three data points which lie close together and have minimal differences in the response

measured. That this corresponds well with CAG's linearization of a single data point with similar dose and response values is also not surprising and is a statement devoid of any interpretative meaning or scientific significance.

c) p. 8, 2nd para. - "The FDA risk estimates from the NCI and Kimbrough data, however, demonstrated remarkable agreement showing that the various strains used reacted similarly to PCB carcinogenic insult."

This sentence because it is so blatantly untrue is perhaps a perfect example of our concerns for the numerous misleading and/or contradictory statements made in the HERD cancer assessment document. The agreement in the risk estimates generated in the Cordell paper, which HERD refers to as the FDA risk estimate, stems from the assumptions made which generated similar response numbers (i.e. total malignancies versus liver cancer) for the two bioassays at the 100 ppm (5 mg/kg/day) dose. Thus, the HERD document clearly, and apparently intentionally, attempts to mislead the uninformed reader into thinking that such manipulation of the data demonstrates that the various strains reacted similarly to PCB carcinogenic insult. The Kimbrough study was a positive carcinogenicity test while the NCI study was concluded to be negative. Two such responses cannot be any more dissimilar in the conclusions reached from the experimental data. That the HERD group quotes the CAG assessment to imply earlier in the document (p.4) that this difference in response was due to the fewer number of animals used in the NCI bioassay represents poor and unfounded speculation on their part. The positive Kimbrough bioassay reported that hepatocellular carcinoma was induced by chronic PCB exposure, the rate for this type of tumor was 14% or 3.5 times the insignificant 4% rate observed for this tumor type in the NCI bioassay. If the two strains had reacted similarly as the HERD document suggests there would have been at least 7

rather than 2 liver cancers observed in the 48 rats exposed to 100 ppm PCBs in the NCI bioassay. A change as large as 5 tumors in 48 animals is not the likely result of having used only 48 animals in this experiment. Besides liver cancer, there were other response differences between the two bioassays. It is this fact that suggests that besides contradicting themselves, the HERD document's mistaken conclusion appears to be intentional for later in the document (p.14 paragraph #2, and p. 18 paragraph #2) the HERD group points out that unlike the Kimbrough study the most frequently found tumors in the NCI study were leukemia and malignant lymphoma. The difference in the responses observed between the two studies when these two tumors are combined for the 100 ppm dose is 0.6% (Kimbrough study) vs. 31% (NCI study) or a 52 fold difference in response. Clearly, the studies yielded totally different responses for various tumor types and came to totally different conclusions which cannot be attributed to a difference in the number of animals used as there is not even suggestive evidence to support such a claim.

Lastly, PCBs are not the only chemical for which the cancer bioassay data generated to date is not consistent, and as such we see no reason to speculate or rationalize in an unfounded manner in an attempt to explain such inconsistencies. If HERD is going to speculate we feel they should do so in a manner consistent with the facts they themselves present and in a way that does not contradict other points they would like to make later in the document. They have clearly committed this error in this instance.

d) p. 14, para. #1 - **"All three risk assessments (CAG, FDA, Crump) are in general poorly documented and justified. It is not always clear whether estimated virtually safe doses or increases in risk apply to for animals or humans."**

We find these statements both amusing and disturbing. Amusing because the HERD group seems to suggest that someone might actually spend their time estimating a 10^{-5} risk for cancer to rats rather than humans. These statements are likewise disturbing, and again we feel quite misleading, as they imply that the other risk assessments including one by the EPA's own Cancer Assessment Group generated "poorly documented and justified" risk assessments. Contrary to HERD we found the CAG's Water Quality Criteria Document to be a much better written and understandable document than the present HERD risk assessment. We would like HERD to document or clearly state where CAG went wrong and to know if the CAG group agrees with the HERD evaluation of their risk assessment. We feel this is a particularly pertinent question as no new scientific information has been presented by the HERD group that suggests that the 1980 CAG assessment is now outdated or inherently wrong. We also find the above HERD statement distressingly misleading to readers for additional reasons. First, it implies that the HERD group knows more about cancer risk extrapolations than either CAG or Dr. Crump because it suggests their work was in error, yet we feel that HERD would not and could not state this directly (particularly as they rely upon Dr. Crump's equation). Second, it implies that the HERD risk assessment is somehow better than the previous three. This is certainly not true. Like many others who have discussed the problems associated with extrapolating cancer risk estimates before, it is appropriate to reiterate that there has been no decision made by the scientific community regarding which model or which calculated virtually safe dose represents the "best risk estimate". This is due to the fact that extrapolations (i.e. 10^{-5} risk) are made down to such low levels of risk that they cannot be measured and verified to test which model is correct or best. Therefore, we are mysti-

fied why HERD feels justified in criticizing other efforts, especially when these efforts do not appear to differ that much from their own estimates of risk. Consistent with these thoughts it is of interest to note that HERD was incapable of "choosing the right model" prior to analysis of the data (see tables 11-16). HERD will only generate confusion and additional questions by the interested public by producing numerous tables demonstrating that depending upon which model they use the risk may differ by 3 orders of magnitude or even greater. Since it is already well known that the model chosen is more important to the final outcome of the risk estimate calculated than is the actual data (OTA, 1981 -see page 163), the HERD statement begs one to ask just what makes their assessment any more valid than the three previous risk assessments? In summary, the above HERD statements are misleading, unsubstantiated, indefensible and in keeping with the overall tone of the document which appears to be a poor attempt to persuade the reader that the HERD assessment by being a newer version is somehow better.

e) p. 21, para. #1 - **"Since NCI stated that under the conditions of this bioassay, Aroclor 1254 was not carcinogenic ... the establishment of a tumor category where significance occurs may seem to be like "fishing" for statistical significance. We do not regard this as a serious objection because there is solid independent confirmation of carcinogenicity from the Kimbrough study ..."**

The HERD division's reference to "fishing" for a finding seems to be a particularly appropriate description of their attempt to lump enough things together until something becomes significant. Perhaps a better adage to paraphase here is "with enough manipulation anything can be made to show a statistically significant difference." Again the HERD division can only rationalize their

actions rather than demonstrate that they acted in a scientifically accepted manner. If the HERD group truly feels justified in their approach then they should explain to the reader the scientific basis for their assumptions and the validity of the manner in which they used the NCI bioassay data (i.e. the scientific precedence for and acceptance of each step in their approach). Along this line it would be particularly helpful to see mention of the regulatory guidelines and the regulatory agencies that perform risk assessments based upon changes in total tumors for negative bioassays. If the HERD division cannot provide such documentation it seems that they have wasted their resources in this effort.

A second and parallel problem with the above HERD statement is that the HERD approach to modeling the data to predict a cancer risk is done in a manner that ignores any understanding of the biological mechanisms by which chemicals might produce cancer. Rather than consider that the mathematical model generating the risk estimates should reflect the carcinogenic process (i.e. the mechanism of the chemical), they feel that any approach is a valid approach to the problem and is somehow scientifically justifiable. We cannot agree. Further, we feel that the risk assessments should be peer reviewed by the scientific community. There is a continuing argument as to which model is "right" and whether or not it reflects the biologic process. A major portion of this debate has no doubt probably centered upon such issues as thresholds in the light of DNA repair, epigenetic mechanisms, time to tumor concepts and other relevant considerations. And while many scientists agree that positive animal data might not correctly identify the human cancer caused, we can find no reason, considering the large toxicological data base on PCBs, to even begin to consider that PCBs are capable of enhancing all forms of cancer.

Consistent with this contention HERD clearly states on page 22 (see comment 3.(d)) that "there is also some question about the validity or advisability of using the category of any malignancy." Therefore, we can only conclude that the HERD risk assessment was not intended to be a scientific attempt at modeling cancer risks nor was it an attempt to reflect the biologic process the mathematical model is supposed to mimic or predict. It is likely that HERD will find little, if any, support for such a simplistic and totally mathematical approach to such a complex problem.

Lastly, the above quote from the HERD document again leads one to ask why HERD even bothered to perform this risk assessment. Much of the text is merely an attempt to rationalize some justification for their "new" and "novel" reassessment of a former EPA effort. Most of the justication HERD puts forth seems to be based upon the fact that when they compare the risk estimates generated by modeling either the Kimbrough or the NCI cancer bioassays, HERD finds it can produce close and similar values. Once again one is tempted to question HERD's assessment, since at virtually safe doses, (i.e. 10^{-5} or 10^{-6} risk) the model is the most important determinant of the risk value generated (see tables 12-16), then what verifiable improvement has the HERD group made. One is also tempted to ask the HERD division whether or not they have determined if lumping all tumors together for the Kimbrough study, as they are doing here, decreases the statistical significance of that study as the PCB exposed group in the study had substantial decreases for several types of tumors.

3. The general text in several places casts doubt upon the experience and competence of the person(s) performing the risk assessment.

a) p. 16, para. #1 - "Liver adenomas are frequently defined to be benign. (Personal conversation with scientists in the Oncology Branch)."

This statement begs the question - How was the risk assessment performed, and by whom? Is it the usual practice of the HERD division to perform a risk assessment using person(s) or staff that don't understand the biologic response they are supposed to be estimating the risks of?

b) p. 21, para. #2 - "However, like adenomas, neoplastic nodules are not malignant and the significance of their appearance is not entirely clear. (Personal conversation with scientists in the oncology branch)".

Same response as above.

c) p. 18, 1st para - "The data for female rats at 1.25 mg/kg/day were also significant; it is not clear whether this is a statistical aberration".

In table 9 it is clear that the 13/24 total malignancies most probably is a statistical aberration as two higher doses have far fewer tumors. If HERD feels that it is not a statistical aberration will they please explain why. The procedure of lumping all tumors together should increase the number of false positives found in any cancer bioassays for which this approach is taken. It seems reasonable to conclude that the 1.25 mg/kg/day for female animals is such a case. Is it not also likely then that the data for the 5 mg/kg/day dose for both sexes likewise represents a "statistical aberration".

d) p. 22, 2nd para - "The NCI study will allow much more mathematically sophisticated risk modeling and measures of goodness-of-fit of the models to the data; but standing alone is difficult to interpret.

Indeed, NCI concluded that it did not establish carcinogenicity. There is also some question about the validity or advisability of using the category of 'any malignancy'".

After admitting that their approach is of questionable validity HERD fails to provide a single scientifically valid reason for their particular mathematical manipulation of the data. This is particularly disturbing in light of the preceding comments as to whether some of the data HERD uses reflects a statistical aberration or a true difference. In table 9 there are numerous sex and dose differences that suggest there was no specific trend for any tumor type with the possible exception of leukemia in the male animals. Since the number in the category modeled, i.e. combined malignancies, reflects a different percent contribution for each tumor type when comparing the doses, how and why does HERD feel justified in combining them? Even after they combine all tumors of a negative study, the NCI data still yields only one positive data point that they feel is statistically different from controls. This is the exact same number as the Kimbrough study (and at the same dose) which HERD felt was inadequate for modeling because it has only one positive data point. Thus, by HERD's own criteria their use of the NCI data is also inadequate.

4. A combination of faulty logic, the inappropriate use of negative data, and exaggerated exposure estimates (see specific comments of Versar's Exposure Estimate document) led to the risk estimates in table 17. The HERD group should re-evaluate their approach and clarify the obvious mistakes in this table because in its present form their estimates of risk are incorrect and therefore are of no practical use.

   a) For inhalation exposures in an occupational setting, table 17 currently boasts a 2.3×10^{-3} mg/yr exposure for an air concentra-

tion of 1.0 mg/m^3. This means that the person exposed only inhales about 2 liters of air the entire year. Yet an average person inhales about 20,000 to 30,000 liters of air/day. Since the EPA's intended value was 1,000,000 times higher in this particular instance, an obvious and glaring mistake, the reader must wonder how many other mistakes table 17 contains.

b) The HERD group risk assessment suggests that highly exposed occupational workers (i.e. capcitor workers) should experience a 2.2 to 3.6 percent excess cancer rate. Such a rate is so high that it should have been easily seen in PCB capacitor workers by now. Since the NIOSH study is negative and because PCBs have been used for over fifty years with no health problems positively attributed to its use other than chloracne, the HERD estimates would appear to represent an obvious gross overestimate.

c) p. 13, 2nd para - "Comparison of results from the FDA, OTA and CAG risk assessment is difficult due to the different units in which risk is expressed in each risk assessment...".

Since the difference in units of previous analyses is a criticism by HERD why didn't HERD generate a 10^{-5} risk in a variety of ways so as to be able to compare the dose that they estimate represents a 10^{-5} risk to those of the previous analyses? Actually, HERD is again misleading the reader with a poor rationalization as they could use the assumptions for fish consumption that were made in the CAG and FDA risk assessments to generate the micrograms of PCBs per day or per year that correspond to a specific risk estimate. In this manner they could compare their values to previous risk estimates. Instead HERD presents their data in a manner that differs from all of the previous

risk assessments thereby compounding a problem they identified with previous risk estimates.

APPENDIX B

HERD's Quantitative Risk Assessment of Reproductive Risks Associated with Polychlorinated Biphenyls

Specific Comments:

1. p. 2, last sentence - **"It incorporates the concept of a threshold dose, because the risk could in fact be zero at any point along the line".**

This meaningless statement is not true. The Gaylor-Kodell technique extrapolates risk along a line connecting zero exposure (i.e. zero risk) and a point representing the upper confidence limit of the lowest data point (which is a non zero value). Therefore, their model cannot have a zero risk at any positive exposure level. It does not incorporate a threshold concept since this would mean the line defining the interpolation would not be estimated through the zero exposure point as it is done in this model but instead the line would reach zero risk at some level of exposure. In its present form this statement appears to be a poor attempt by the HERD group to rationalize their approach and misleads the reader by stating the model reflects an important concept in risk assessment which the model does not.

2. p. 4, 2nd para - **"As the rat study did not have the problems (mentioned previously) of the Rhesus monkey study the rat data was selected for the risk assessment".**

If the monkey data was inadequate for various reasons as is stated here in the executive summary then why is so much of this document spent discussing this data? Why wasn't a simpler document written using the rat data only?

3. p. 4, last para - "The scenarious for which the risk is greater than 1 in 100,000 are 1) exposure at the level of quantitation for PCBs in air 10 mg/m^3; ...".

A level of 10 mg/m^3 is a very high air concentration of PCBs and not a realistic consideration as it is 10 times the allowable level for the workplace as regulated by OSHA. As such it is a level so high it is not reasonably expected to be produced anywhere in the United States at the present time and certainly not for a daily exposure during a 40 year exposure period.

4. p. 5 through 15 - Contains numerous statements concerning the monkey studies that undermine their usefulness for risk assessment purposes. Again one wonders why this data was discussed by HERD and to what purpose it serves. The following are cited examples:

   a) p. 6, 1st sentence - "Generally, the study protocols has been reported in a confusing manner in several articles by Allen et al.".

   b) "The analyses of the adipose tissue levels of PCBs in mothers which had presumably not been dosed with PCBs for one year were variable and probably near the limit of accurate measurement. The errors in the estimation of these values were between ± 74% and ± 112%".

   c) p. 9, 3rd para - "She stated that the rhesus monkey has been found to be more sensitive to PCBs than other species of nonhuman primates. The reason for this sensitivity is unknown".

   d) p. 10, 1st para - "Some of the variation was due to the method of analysis".

e) p. 11, 2nd para - "It is also possible that large errors were made in the determination of tissue PCB levels".

f) p. 11, 2nd para - "Methods which use an extraction procedure, a chromatographic clean-up and gas chromatogram quantification of multicomponent systems as was done in these studies have larger errors associated with the data obtained by them".

g) p. 11, 3rd para - "Another potentially large error is inadvertant contamination...". or p. 12, 1st para - Since the amount causing toxicity is apparently so small one cannot eliminate the possibility of contamination".

h) p. 15, ast para - "As mentioned previously, Rhesus monkeys may be much more sensitive to PCBs than humans. Also, at some of the higher dose levels, mothers were showing signs of PCB toxicity (unfortunately, there are no data on maternal weight); thus it is possible that the lowered birth weights of an infant may have been due to the mother's restricted intake rather than toxicity in the offspring".

5. p. 11, 1st para - Thus, 71.3 divided by 2 is approximately equal to 36. Therefore, these PCBs in female monkeys not dosed for one year are more than 36 times more toxic to infant monkeys than the PCBs administered to the mothers".

This statement is not a scientifically valid conclusion. The level of PCBs in the mothers not dosed for one year does not give the HERD group any estimate of the level of exposure of the infant monkeys nor their PCB body burdens. Thus, HERD has no way of knowing from this data whether or not the PCBs remaining are more toxic and if so by how much. A perusal of some

of the Allen et al. papers indicates that amongst infants born to PCB exposed monkeys there is a great degree of individual variation in the tissue levels of PCBs. This variability makes generalizations about tissue levels difficult. In addition, the paucity of quantitative body burden data does not allow one to clearly discriminate whether or not infant Rhesus monkeys are inherently more sensitive to a given dose of PCBs when compared to adults. Thus, not only is HERD's conclusion of this calculation specious reasoning but there is not sufficient data in the Allen et al. reports to make such an estimation.

6. p. 14, 1st para - **"This exercise obviously has its limitations. For example, if the maternal dose or exposure in mg/kg/day is very high then the 12 week body weights will become very small or even negative".**

After listing numerous limitations of the monkey data, HERD then generates an equation based upon this questionable data base which is of even more questionable utility. Since birth weights cannot be negative and as infant mortality should in all likelihood occur before some lesser but positive value is reached, for what dosage range of PCBs and infant birth weights might this equation be appropriate? Since HERD has no idea concerning what usefulness this equation might serve, nor was it utilized in the final risk assessment, why was this equation even included in this document? More to the point, since female capacitor workers have been studied, a preliminary report of which was presented at a PCB Symposium in 1982, why is the HERD review group only relying on monkey and rat data?

7. p. 19, 1st para - **"Another way of estimating 'safe' doses for exposures to PCBs is by modeling of the type used for quantitative risk**

assessments for cancer. Many toxicologists support only the NOEL approach because they believe that the alternative of modeling teratogenic or reproductive effects implies that there is no threshold for those effects. This belief is mistaken, however, because actual or virtual threshold can be accomodated with certain extrapolation models".

Since HERD does not use a model which estimates a threshold (the Gaylor and Kodel modal certainly does not), perhaps they could explain how their approach alleviates the concern "many toxicologists" might have for their modeling of the data. We doubt that they can. Furthermore, HERD is taking great liberties in describing why certain toxicologists might prefer the NOEL approach and apparently inserts their own unsupported opinion for that of others. Did HERD poll the scientific community before deciding why many toxicologists support a NOEL approach?

8. p 19, 1st para - "In addition, there is no convincing reason to believe that thresholds should occur for all chemicals with all kinds of reproductive effects or all human populations".

HERD ignores the fact that there is likewise no convincing scientific reason to support the claim that no chemicals have thresholds for reproductive toxicities. That HERD attempts to justify their approach in this manner represents poor and unsupportable scientific reasoning.

9. p. 19, 3nd para to 20, 1st para - This should be re-written as there are numerous errors in this paragraph and it seems certain that wnatever point the author is trying to make will escape the reader. For example on page 20 - "Even when four events occur in the treated group, with none in the controls, results will no reach the usual 50%

significance level if each sample is 17 or more. In most instances the sample size should be chosen to be capable of producing at least five events in the treated group when none occur among controls, when hazards of concern are persent".

The scientific community at large would probably be most interested in learning how the HERD group picks the appropriate sample size that will yield five or more responses for a toxicity test with no prior knowledge of the dose response curve.

10. (p. 20-21 - Concerning the quotation from a 1981 CEQ document). The quotation that HERD reprints here does not state that thresholds do not exist and therefore does not necessarily support their approach.

11. p. 27, 2nd para - "Since the F2a generation in Sherman rats has four dose levels plus control, the best picture of dose response can be obtained from that study. Thus the F2a generation in Sherman rats is used as the basis for extrapolating risk to humans at environmental, occupational, and consumer exposure levels".

The HERD group is misusing the linear interpolation technique and the data. Only the highest dose appears to provide an adverse response (i.e. statistically different from control), so that using all of the data points to generate confidence limits for a dose response that doesn't exist represents poor science. To claim that a dose response relationship exists is an inaccurate description of a single positive data point. Again, modeling negative data points only generates meaningless numbers. In addition, since much of the data the HERD group is using is not significant it suggests that a threshold exists. Yet, contrary to the data, HERD uses a linear nonthreshold model to generate its risk numbers. If HERD is going

to generate such meaningless numbers to what credible use can such numbers be put? We feel there is none.

12. p 27, last sentence - "It (linear interpolation) is also generally the most conservative model and thus the model most protective of public health. Thus, it was the model used to extrapolate risks to humans". The HERD group attempts here to fool themselves and the reader into believing that they have in fact used a mathematical model to determine risk. In actuality they have not for as HERD so aptly stated on p. 19 there is no difference in using a safety factor, i.e. a NOEL approach, or linear interpolation. HERD's 10^{-5} risk estimate is derived in actuality by using a safety factor of 100,000 on the upper confidence limit of the expected response for some dose. Although HERD might prefer to say that such an approach is mathematical modeling, it is the simple application of a larger safety factor which does not improve the reliability of their approach. As the saying goes "a rose by any other name is still a rose". Thus, if HERD feels that 1,000 is too small a safety factor then let them state it clearly, rather than attempting to obscure the outcome by insisting their numbers were generated by a "better model".

HERD stated on p. 19 that there was no reason to believe that all chemicals possess thresholds, likewise there is no convincing evidence that suggests a dose 1/100,000 of that producing an adverse response yields 1/100,000 of the given response, or that it produces an equivalent response once in a hundred thousand exposures. Herein lies the fallacy of HERD's numbers. HERD has adopted the simplistic but unrealistic interpretation that for reproductive toxicity as the exposure is lowered the occurence of the adverse effect becomes less likely while the magnitude of the adverse

response induced remains unchanged. This ignores the traditional concept that there is a spectrum of toxic effects for any chemical for which the severity decreases with the dose. Moreover, as HERD states elsewhere, the numbers do not represent the actual risk but an estimate of the upper bound of the risk. That is, the actual risk for a given exposure is less than the risk given by linear interpolation. Since the HERD approach has all of the above flaws associated with its numbers, and since they do not know how much less the actual or real risk is for a given dose, then what use are their numbers? This is a particularly relevent question as the actual risk may be several magnitudes lower just as their reproductive risk estimates varied by several magnitudes depending upon which model is used (see tables 8, 9, 10 etc.).

13. Table 13 - **"Reproductive Risk of Death Prior to Weaning of Human Infants From Hypothetical Maximum Exposures of Pregnant Women to PCBs (Derived Using the Technique of Linear Interpolation)".**

This table which is a summation or culmination of HERD's reproductive risk assessment contains numerous inconsistencies and apparent mistakes. HERD is first urged to re-evaluate this risk assessment, then correct their typographical mistakes and publish a better discussion concerning what these values represent as well as the credibility or uncertainty inherent to these values. For example:

a) On the first page of table 13 one reference scenario poses the improbable workroom air concentration for PCBs of 10 mg/m^3. This air concentration is unrealistically high and probably cannot be shown to exist anywhere in the United States. Yet a cross check of the Versar document shows this value to be a typographical

error both in table 13 and on page 5 of the Executive Summary. How many other typographical or mathematical errors exist in this document and have become entrenched in its conclusions?

b) For each estimate of reproductive risk the flaws or uncertainties linked to the methods and assumptions used are combined with the uncertainties and inaccuracies inherent to the exposure estimates produced by Versar that were then used to determine the level of risk as set by the "model". There are numerous unrealistic assumptions made in the Versar document (see comment (c) below). Since the estimates derived by the model using such exposure data can be no more valid or accurate than the data, much if not all of the risk estimates are no more than meaningless assumptions. It is suggested that the exposure estimates be re-calculated in a more believable, accurate, and realistic manner before HERD even attempts to propose doing any reproductive risk estimates by any model.

c) The risk estimate for plastic building materials is only 1/10 of that for a constant air exposure of 10 ug/m^3. This essentially means that the room air concentration of PCBs caused by the plastic is 1.0 ug/m^3. Considering the low vapor pressure of PCBs, the fact that it is part of a plastic matrix at low concentrations and unlikely to move to the surface, the exposure value proposed by Versar and the subsequent risk estimated by HERD do not appear to be even remotely possible. (Note: p 259 of the Versar document indicates that exposure estimates used here by HERD are indeed only hypothetical). Therefore, EPA should re-

consider the exposure estimates used, in this risk assessment and should directly measure the actual PCB levels before it makes any decisions of a regulatory nature.

d) Since very few of the HERD calculated risks exceeded the 10^{-5} risk level and as these tended to be generated by spurious exposure estimates, the HERD document should state in the executive summary that their calculations indicate that there is no apparent reproductive risk associated with incidentally produced PCBs.

e) A comparison of the risks of cancer or reproductive toxicity using the upper 95% confidence limits for both values reveals that the HERD group is suggesting that the reproductive risk is slightly greater than the cancer risk at all exposure levels. We challenge the HERD division to propose the mechanism of reproductive toxicity and the data that is consistent with this type of risk assessment modeling particularly for low level exposures. (It should be pointed out that the "positive" data point was a 100 ppm dose for both the cancer and reproductive risk assessments. This highlights the "model dependent" nature of their assessments). Not only are we certain that they cannot, we feel likewise certain that the scientific community would not agree with such speculation, nor will they agree with the type or risk assessment performed here by HERD. The inconsistency of HERD's approach to modeling oncogenic and reproductive toxicities only serves to reinforce the conclusion that HERD can offer no scientific evidence to support their calculations. HERD states on p.

30 that the purpose of this risk assessment is to provide guidance in establishing a permissible level of PCBs in other chemicals. But since their risk assessment is obviously flawed it is concluded that these documents should not be relied upon for regulatory guidance.

APPENDIX C

VERSAR'S EXPOSURE ASSESSMENT FOR

INCIDENTALLY PRODUCED POLYCHLORINATED BYPHENYLS (PCB)

The following citations from the Versar document indicate that many of the exposure estimates either are based upon unjustified or simplistic assumptions that do not reflect actual or real world conditions for PCB exposures. This is not an exhaustive listing but only a few cited examples are provided to demonstrate our concerns for the accuracy of the exposure estimates used by HERD as these in turn helped establish risk estimates for certain scenarios.

1. p. 40-42 - Using simplistic assumptions for only two chlorinated biphenyls, it is calculated that a highly chlorinated isomer would not enter the groundwater. However, this value is ignored and an amount representing the monochlorobiphenyl contamination estimate, a chemical whose environmental transport and fate is distinctly different from PCB mixtures, is used for exposure purposes. When combined with the rest of the worst case assumptions made in this scenario the estimates have even less practical use.
2. p. 53 and table E-1 on page 54 - PCBs are assumed to be present in the workroom air at the same ratio to the process chemical as is found in the production stream. Considering the extremely low vapor pressure of PCBs this could not be true and greatly overestimates the percent of PCB present in the workroom air.

3. p. 88 - An assumption is made that the only process affecting removal of PCBs from the water is volatilization and not absorption to the sediments. Versar admits that this is in fact a scenario contrary to reality but proceeds anyway.

4. p. 104 - Versar assumes a level of PCBs in grain that exceeds that found in grain from FDA sampling surveys. Further, PCBs were rarely found in grains sampled by the FDA. Thus, the Versar estimate does not attempt to reflect reality.

5. p. 216 - The Versar document states "the effect of duration (for soap exposure) is not taken into account when 100% absorption is assumed ... even though soaps are generally washed off within minutes of application."

6. p. 253 - The release rate of PCBs from dried paint is apparently estimated as greater than the evaporation rate of PCBs from a liquid coating. In addition, all of the PCBs is assumed to be released in two years rather than most of it remaining in the paint which is the more plausible prediction.

7. p. 259 - While there is no evidence that PCBs occur in polyvinyl chloride plastic, the Versar document hypothesizes a significant PCB exposure from such plastics.