SAP Report No. 99-05
November 18, 1999

REPORT

FIFRA Scientific Advisory Panel Meeting, September 21-24, 1999, held at the Sheraton Crystal City Hotel and Days Inn Crystal City Hotel, Arlington, Virginia

Sets of Scientific Issues Being Considered by the Environmental Protection Agency Regarding:

Session I - Issues Pertaining to the Assessment of Residential Exposure to Pesticides
Session II - Review of An Aggregate Exposure Assessment Tool
Session III - Identifying Carbamate Pesticides that Have a Common Mechanism of Toxicity
Session IV - Proposed Guidance for Conducting Cumulative Hazard Assessments for Pesticides that Have a Common Mechanism of Toxicity
Session V - Review of American Cyanamid Company's Probabilistic Assessment for Chlorfenapyr and Request for Guidance on Problem Formulation

NOTICE

This report has been written as part of the activities of the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Scientific Advisory Panel (SAP). This report has not been reviewed for approval by the United States Environmental Protection Agency (Agency) and, hence, the contents of this report do not necessarily represent the views and policies of the Agency, nor of other agencies in the Executive Branch of the Federal government, nor does mention of trade names or commercial products constitute a recommendation for use.

The SAP was established under the provisions of FIFRA, as amended by the Food Quality Protection Act (FQPA) of 1996, to provide advice, information, and recommendations to the EPA Administrator on pesticides and pesticide-related issues regarding the impact of regulatory actions on health and the environment. The Panel serves as the primary scientific peer review mechanism of the EPA, Office of Pesticide Programs (OPP) and is structured to provide balanced expert assessment of pesticide and pesticide-related matters facing the Agency. Food Quality Protection Act Science Review Board members serve the SAP on an ad-hoc basis to assist in reviews conducted by the SAP. Further information about SAP reports and activities can be obtained from its website at http://www.epa.gov/scipoly/sap/ or the OPP Docket at (703) 305-5805. Interested persons are invited to contact Larry Dorsey, SAP Executive Secretary, via e-mail at dorsey.larry@.epa.gov.

TABLE OF CONTENTS

SAP Report No. 99-05A, November 18, 1999

REPORT:
FIFRA Scientific Advisory Panel Meeting, September 21, 1999, held at the Sheraton Crystal City Hotel, Arlington, Virginia

Session I - A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding:

Issues Pertaining to the Assessment of Residential Exposure to Pesticides

Mr. Larry Dorsey
Designated Federal Official
FIFRA/Scientific Advisory Panel
Date:_____

Ronald J. Kendall, Ph.D
Chair
FIFRA/Scientific Advisory Panel
Date:_____

Federal Insecticide, Fungicide, and Rodenticide Act
Scientific Advisory Panel Meeting
September 21, 1999

SESSION I - Issues Pertaining to the Assessment of Residential Exposure to Pesticides

PARTICIPANTS

Chair
Ronald J. Kendall, Ph.D., Professor and Director, The Institute of Environmental and Human Health, Texas Tech University/Texas Tech University Health Sciences Center, Lubbock, TX

FIFRA Scientific Advisory Panel
Charles C. Capen, D.V.M., Professor and Chairman, Department of Veterinary Biosciences The Ohio State University, Columbus, Ohio

FQPA Science Review Board Members
Ross B. Leidy, Ph.D., Department of Toxicology, North Carolina State University, Raleigh, NC
Natalie C. Freeman, Ph.D., Department of Environmental/Community Medicine, Robert Wood Johnson School of Medicine, Piscataway, NJ
John Kissel, Ph.D., P.E., Department of Environmental Health, University of Washington, Seattle, WA
Andrew Smith, S.M., Sc.D., Maine Bureau of Health/Environmental Toxicology Program, Augusta, ME
Katherine Shea, M.D., M.P.H., Chapel Hill, NC
Mr. Gerry Akland, RTI, Knightdale, NC
Mark Miller, M.D., M.P.H., Chico, CA

Designated Federal Official
Mr. Larry Dorsey, FIFRA Scientific Advisory Panel, Office of Prevention, Pesticides and Toxic Substances, Environmental Protection Agency, Washington, DC

PUBLIC COMMENTERS

Oral statements were made by:
Dr. Muhilan Pandian, American Crop Protection Association
Ms. Leah Rosenheck, Outdoor Residential Exposure Task Force
Dr. John Deprosto, Outdoor Residential Exposure Task Force
Dr. Susan Youngren, Responsible Industry for a Sound Environment
Dr. Dave Esterly, Spray Drift Task Force
Ms. Usha Vedula, Indoor Residential Exposure Joint Venture
Mr. William Perlberg, Hartz Mountain Corporation
Dr. Jim Clark, BASF Corporation
Dr. David Wallinga, Natural Resources Defense Council
Dr. Larry Smith

Written statements were received from:
Dr. Muhilan Pandian, American Crop Protection Association
Ms. Leah Rosenheck, Outdoor Residential Exposure Task Force
Dr. John Deprosto, Outdoor Residential Exposure Task Force
Dr. Susan Youngren, Responsible Industry for a Sound Environment
Dr. Dave Esterly, Spray Drift Task Force
Ms. Usha Vedula, Indoor Residential Exposure Joint Venture
Mr. William Perlberg, Hartz Mountain Corporation

INTRODUCTION

The Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Scientific Advisory Panel (SAP) has completed its review of the set of scientific issues being considered by the Agency regarding issues pertaining to the assessment of residential exposure to pesticides. Advance notice of the meeting was published in the *Federal Register* on September 3, 1999. The review was conducted in an open Panel meeting held in Arlington, Virginia, on September 21, 1999. The meeting was chaired by Ronald J. Kendall, Ph.D, Professor and Director, The Institute of Environmental and Human Health, Texas Tech University/Texas Tech University Health Sciences Center, Lubbock, TX. Mr. Larry Dorsey served as the Designated Federal Official.

The 1996 Food Quality Protection Act (FQPA) requires the Agency to routinely address nondietary and non-occupational pesticide exposure for the general population. These are exposures that can occur in a residential setting (or other areas frequented by the general population) and that do not occur as part the diet or as a result of participation in occupational practices. In response to FQPA, the Agency developed Standard Operating Procedures (SOPs) for residential exposure assessment, which it brought before the SAP on September 9, 1997. Today's meeting does not present a revised version of the 1997 SOPs. Instead, the Office of Pesticide Programs (OPP) is presenting the most critical issues for discussion prior to developing a revised SOP document. These critical issues are: calculating percent dislodgeability of available pesticide residues from lawns, indoor surfaces, and pets; use of choreographed activities as surrogates for estimating children's dermal exposure; characterizing hand (or object)-to mouth activities; calculating exposure to pesticides that may result from track-in, spray drift, bathing or showering; estimating exposure of children of farmers or farm workers to pesticides; exposure to drift; and calculating exposure from use of pesticides in schools, day-care center, and other public places. OPP requests the Panel's input on these issues and responses to specific questions concerning these exposure issues.

CHARGE

The specific issues to be addressed by the Panel are keyed to the Agency's background document, ***Overview of Issues Related to the Standard Operating Procedures for Residential Exposure Assessment***, dated August 5, 1999 and are presented as follows:

Issue #1 - Percent Transferable Residues

1. OPP is proposing to change the default assumptions in its SOPs for "percent transferable residues" of pesticides on lawns, indoor surfaces and pets. Does the Panel find these changes reasonable and scientifically defensible, based upon the available data? In particular, does the Panel agree with OPP's proposed assumption of 5% transferability for indoor surfaces, recognizing that data for carpet and desktops support this level, but data for vinyl surfaces show 10% to 20% transferability? Similarly, should OPP consider using a higher "percent transferable residue" factor for wet surfaces and/or sticky hands or not?

Issue #2 - Surrogates for Estimating Dermal Exposure to Children

2. OPP has indicated the intention to continue to use choreographed activities by adults as surrogates for estimating dermal exposure to children. Specifically, OPP has proposed the use of 20 minutes of Jazzercise as a surrogate for up to 4 hours of mixed activities. This position is based on comparisons to biological monitoring studies with adults performing choreographed activities. The Panel is asked to comment upon this approach and its utility when addressing short-term exposures (1 - 7 days) or exposures of longer durations. In addition, the SOPs currently do not account for potential differences in permeability of children's skin compared to adult skin and the Agency has found no scientific data to document such differences. How does the Panel think that the SOPs should address the concern that infants' and children's skin may absorb pesticides at a greater rate than adult skin?

Issue #3 - Frequency of Events

3. OPP has adopted the SAP's previous recommendations concerning the frequency of hand-to-mouth events (20/hr) and available hand surface area (20 cm^2). Are these assumptions protective of teething toddlers (8-18 months old), particularly concerning the amount of the hand placed in the mouth (two to three fingers; 20 cm^2)? The frequency of 20 events per hour is the 90^{th} percentile from a study involving observations of children at home and in day care centers. The mean in that study is ~10 events per hour. Panel is also asked to comment on the use of these values when addressing short-term exposures (1 - 7 days) or exposures of longer durations.

Issue #4 -Estimating Exposures from Secondary Sources

4. Given the relatively low magnitude of exposures from track-in, bathing or showering relative to other scenarios, should OPP estimate exposure to pesticides that may result from these sources? If so, have we identified the most critical scenarios and approaches to be used to do the estimation?

Issue #5 - Estimating Exposures from Non-residential Pathways

5. OPP proposes to address exposure of children living on or near farms where pesticides are used by estimating deposition on lawns resulting from pesticide drift; OPP is developing a drift model for this purpose. Does the Panel consider this approach reasonable and are there other important non-residential pathways of potential pesticide exposure that should be evaluated for farm children?

Issue #6 - Addressing Exposure from Spray Drift

6. OPP is proposing to initiate the use of a spray drift model to estimate the likely magnitude of unintentional exposure to pesticide residues as a result of direct exposure to sprays. What is the Panel's opinion concerning the introduction of this new source of exposure into the risk assessment process?

Issue #7 - Twenty- four Hour Assumption Used in Estimating Risk in Schools, Day Care Centers, and Other Public Places

7. OPP currently assumes 24 hour residential exposure as a basis for its exposure assessments. OPP believes that this assumption is sufficiently conservative to protect from exposures that are likely to be encountered in other non-residential settings such as schools, day care centers, or other public places where the use patterns are comparable. Does the Panel agree or disagree and why?

PANEL RECOMMENDATION

Medically, a screening tool is designed to be highly sensitive (e.g., few false negatives) often requiring a trade-off in being less specific (e.g., allowing more false positives). If the Standard Operating Procedures (SOPs) are to be used as a screening tool, they should reflect this orientation and choices should err on the side of overestimating exposures. Thus, using means and other measures of central tendency would not be appropriate. Rather choosing "numbers" that reflect the right side of all distributions, be it the upper limits of the range of measurements when few data are available or the upper bound of a 95^{th} or 99^{th} percentile, is much more conservative and protective.

DETAILED RESPONSE TO THE CHARGE

Issue #1 - Percent Transferable Residues

1. OPP is proposing to change the default assumptions in its SOPs for "percent transferable residues" of pesticides on lawns, indoor surfaces and pets. Does the Panel find these changes reasonable and scientifically defensible, based upon the available data? In particular, does the Panel agree with OPP's proposed assumption of 5% transferability for indoor surfaces, recognizing that data for carpet and desktops support this level, but data for vinyl surfaces show 10% to 20% transferability? Similarly, should OPP consider using a higher "percent transferable residue" factor for wet surfaces and/or sticky hands or not?

In general, additional research is required to develop a realistic percentage of pesticide transfer from different surface types. A survey of peer-reviewed articles should be conducted to determine a realistic range of transfer rates. In addition, formulation types (e.g., microencapsulated) should be evaluated and compared (e.g., microencapsulated versus emulsifiable concentrate). Until such data have been evaluated or studies conducted to determine transfer rates from specific surfaces, the Agency should consider using a more conservative approach, (i.e., 20% from surfaces that have no supporting data).

The need exists for more discussion about studies validating methods used to estimate transferable residues from hard surfaces, such as floors, counters, decks, etc. Of fundamental concern is whether methods show a linear relationship between pesticide removal and the area wiped. If the relationship is not linear, comparison of wipe data across different studies employing different sampling designs becomes problematic.

(1) ***Lawns***: Although based on most of the current literature cited in the report, the 5%

transferable residue is most likely an overestimation. Furthermore, studies by Cowell et al. (1993) and Hurto and Prinster, (1993) for two separate pesticides suggest that 5% may not provide a suitable margin of error for all pesticides potentially applied outdoors. In these two cases, the percent transferable residue was 4% for isofenfos, and for Dithiopyr-Microencapsulated the dislodgeable residues were 3.19%. The Agency should determine what would be a suitable margin of error in this case. The microencapsulated pesticide raises another interesting issue that the Agency needs to address, because the formulation of the pesticide may play an important role in its transferability. A microencapsulated formulation could have a greater attachment potential, because it "sticks" to an insect's body and is ingested by grooming.

However, one Panel member expressed that the 5% transfer rate was too low, as more material could be transferred from hard surfaces and toys. There was concern that variation between a hard surface (e.g., floor) and clothing may not be equal. A member of the public commented that the use of a simulated dermal press removed all pesticide residues (compound not named) from plastic but only 4% from carpet. A Panel member asked if different carpet types had been tested to determine removal efficiency. It was stated that a few types had been examined, but the primary material was nylon.

A Panel member questioned if there were data on validating the removal of pesticides from surfaces by the use of repetitive motions for X number of times on the same area. The Agency stated that there were limited data available on validations. It was pointed out that wetness of a surface, human or turf, could have an effect on transferability of residues.

One Panel member questioned the rationale for choosing the PUF roller as the standard for surface transfer residuals since the method seems to affect the values greatly. For example, the shoe method may be more realistic than the roller for activities like playing soccer or football on a recently treated field. Having the rationale fully explained along with biomonitoring data showing the correlations with equivalent surfaces tested with the various methods coupled with urinary metabolite data would be very useful. Even if the PUF roller method is more routinely used and easier to standardize, if it does not accurately reflect the transfer residuals when compared to biomonitoring data, then it is not the best method to choose.

(2) ***Indoor Surfaces***: Why does the Agency assume uniformity in "percent transferable residues" for all indoor surfaces? The validity of the proposed change depends upon the relative frequency of touches involving hard and soft surfaces. What defaults are built into Stochastic Human Exposure and Dose Simulation (SHEDS)? Shouldn't the SOPs and SHEDS reflect similar (if not exactly equivalent) assumptions? The 5% transfer rate for indoor surfaces would be inappropriate based on current literature. Both Camann (1995) and Fenske (1990) found substantially higher transferable residues (23.5% and 11.4%, respectively). Homes contain both smooth and textured surfaces. Since the range of surfaces available in homes are variable, the more conservative approach would be to accept the higher residue transfer values. A question was asked if different types of carpet had an effect on surface residue levels, and it was stated that no data were available to address this concern. More data are needed from surfaces other than carpet.

(a) In support of a more conservative value, it should be noted that the median contact rates by 30 preschool children with smooth surfaces was 80 times per hour; while for textured surfaces such as carpet, the median contact rate was only 16 times per hour (Reed et al. 1999). In addition, the Minnesota Children's Pesticide Exposure Study observed 19 older children (ages 3-12 years) and showed a similar directional difference (Freeman et al. 1999). The greater contact with surfaces with higher transferability should mean that a more conservative figure (i.e., a higher transfer rate) be used rather than a less conservative figure.

(b) Children often have wet and sticky hands, feet, faces, abdomens and chests from saliva. Moist and sticky hands of children seem to be an important concern. Saliva extracts residue with higher efficiency in some studies for some chemicals. They can also be wet in the groin/buttocks area due to urine, or all over the body due to activities or other bodily secretions. This potentially affects non-dietary ingestion as well as dermal absorption. With respect to moist or sticky hands, there are not enough available data to make a determination whether using a higher "percent transferable residue" factor is justifiable. Therefore, it would be better for the Agency to err on the side of the higher transfer rate until further data are available.

(3) ***Pet applications***: We agree that more research is needed with respect to pet applications. Thus far, the transfer models are based on the assumption that the major route of exposure is from petting the animal followed by licking or mouthing the hand. The Agency should also consider that children kiss, mouth, lick, cuddle and sleep with their pets, as well as handle and eat food with hands that have just contacted the pet. The comment was made that it may be safe to assume that children would not play with animals the first 24 hours after they are treated. This is counter-intuitive, because within hours after treatment would be the most pleasant time to play with a clean, fluffy, nice smelling, flea free dog. Therefore, this is not a justifiable assumption.

One Panel member indicated that choosing the mean of 20% for the pet transfer (as stated on page 39 of the Agency's background document) is not consistent with the concept of screening, where central tendencies should be avoided and cut offs at the upper bounds should be chosen.

Issue #2 - Surrogates for Estimating Dermal Exposure to Children

2. OPP has indicated the intention to continue to use choreographed activities by adults as surrogates for estimating dermal exposure to children. Specifically, OPP has proposed the use of 20 minutes of Jazzercise as a surrogate for up to 4 hours of mixed activities. This position is based on comparisons to biological monitoring studies with adults performing choreographed activities. The Panel is asked to comment upon this approach and its utility when addressing short-term exposures (1 - 7 days) or exposures of longer duration.

In addition, the SOP's currently do not account for potential differences in

permeability of children's skin compared to adult skin and the Agency has found no scientific data to document such differences. How does the Panel think that the SOP's should address the concern that infants' and children's skin may absorb pesticides at a greater rate than adult skin?

A comparative study, using children and adults in similar activities, is required to assess dermal exposure in children using the adults as surrogates. Children's behavior patterns (i.e., aggressive vs. passive behavior) must be taken into account when determining potential exposures. The Agency must consider children's ages when assessing dermal exposure, because the literature describes variability when age is used as a criterion in the evaluation process. Future studies involving school children must include the collection of like samples (e.g., urine) from their parents in order to determine if the amount of exposure is greater than other age groups in the study.

It is illogical for the Agency to assume that short, vigorous activity is equivalent in exposure potential to longer periods of more passive behavior. Without adequate evaluation of the adult's Jazzercize behavior relative to activities actually conducted by children (i.e., the types, frequency and duration of contacts in a twenty minute period, and concurrent inhalation and cardiovascular rates), it is unclear that adult Jazzercise tells much, other than what an adult is exposed to during a 20 minute routine. Some sort of comparative study is needed to assure that the extrapolations proposed by the Agency are reasonable. The proposed conversions are based on adult-adult not adult-child comparisons. Some of the adult surrogate activities compared to Jazzercise (e.g., picnicking on a blanket) appear to mimic very passive behavior. An approach that took into account the relative frequency (among children) of aggressive and passive behavior would be more convincing than the flat assumption of equivalence without examination of behavioral patterns.

Comments were made suggesting the use of a probabilistic approach to determine dermal exposures by looking at both aggressive and passive activities. It was pointed out that industry has dermal exposure studies which it is sharing with the Agency. A question was raised on both bioavailability and persistence of a pesticide over time and how this would affect transfer to a body surface. No specific answer was given.

Some of the logic presented by the Agency concerning lawn contact is flawed. There is very little contact with grass as it is being mowed. In contrast, sitting or lying on grass provides a longer period of contact. Perspiration on the limbs in contact with the grass could facilitate transfer.

It should be noted that studies can be conducted with children in "real" environments where parents routinely use pesticides without raising the ethical issues that concern the Agency. It should also be noted that there are always logistical problems when doing studies with people in homes (regardless of age) in comparison to using employees in a laboratory. The laboratory studies, such as Jazzercise provide only an initial approximation. Without real world validation, it is difficult to interpret the Jazzercise results. There are post-application studies being conducted now with children in real world situations that might illuminate these issues, but the results will

not be available for some time.

One Panel member raised the question, "are weight and surface area the appropriate scaling factors for back extrapolating Jazzercise results to children"?

Of particular interest in the reported Jazzercise data, the participants' exposures are as great or greater at 9 hours after application as they were 3 hours after application. Since the approved re-entry time post-application is typically 1-2 hours, what does this say about reasonableness of current re-entry standards? The study by Shah (1987) which the Agency uses to argue that there is no “major” difference in permeability between children’s skin compared to adult skin does not support the Agency's argument of no "major" difference in skin permeability. Some of the pesticides showed higher uptake in younger animals, and others did not, while still others showed higher uptake in the older animals. It is believed that this shows that we need to learn more about how the various pesticides act and that it may not be appropriate to try to make a “one size fits all” dermal exposure model for pesticides.

It should be noted that the Shah study used 33 day-old rats for its young group. This is akin to using a pre-pubertal adolescent as a surrogate for an infant or toddler. One would not expect many differences between a nearly mature rat of 33 days and an adult rat of 84 days; thus, the fact that differences were found is very interesting. In addition, pediatricians are taught that children absorb more through the skin than adults because of the increased surface area/volume ratio as well as differences in skin permeability. The Agency appears to be satisfied that these differences are related to sampling technique, but the primary data were not supplied. Thus, this judgement was made after a conversation with the authors. Explicitly, further discussion of this point is warranted, because it goes against “established wisdom.” One also wonders if a rat is an appropriate dermal model for a human.

The Minnesota Children’s Pesticide Exposure Study conducted in collaboration between the University of Minnesota and the RTI/EOHSI/NHEXAS Consortia has produced some very interesting pesticide exposure information. When compared to the NHANES chlorpyrifos metabolite data, the children in the NHEXAS study had higher metabolite levels. Unfortunately, urine samples were not collected from adults in the same NHEXAS families for comparison with their children. The chlorpyrifos loading on the hands of the NHEXAS children did not differ across ages; however, there was a slight but significant negative correlation between the age of the child and metabolite levels suggesting that the dermal absorption was greater in the younger children or that the younger children’s activity patterns increased their exposure relative to older children.

It was stated that it would be useful to have access to the studies related to children’s exposure on the Internet. This would allow the Panel members the opportunity to review the primary data and foster judgements about the choices entertained by the Agency, rather than requiring the Panel members to rely either on summary data, or search for articles independently.

Issue #3 - Frequency of Events

3. OPP has adopted the SAP's previous recommendations concerning the frequency of hand-to-mouth events (20/hr) and available hand surface area (20 cm^2). Are these assumptions protective of teething toddlers (8-18 months old), particularly concerning the amount of the hand placed in the mouth (two to three fingers; 20 cm^2)?
The frequency of 20 events per hour is the 90th percentile from a study involving observations of children at home and in day care centers. The mean in that study is ~10 events per hour. The Panel is also asked to comment on the use of these values when addressing short-term exposures (1 - 7 days) or exposures of longer duration.

Additional data are required in order to determine the frequency of hand-to-mouth (also hand-to-feet in infants) contact among age groups. Since the sample size is so small from the few studies published, the upper range of the distribution from these studies should be used when determining absorption from hand-to-mouth activities. Although 20 hand-to-mouth events appear to be an appropriate measure over a one- to six-day period, it is not known what the effects of developmental changes are having on this value. Also, the dorsal surface of babies' hands must be taken into account, because this portion of the hand is used when on flat surfaces. Additional studies are underway in Arizona and Texas, which might provide some insight into this important aspect of research. To the extent current data on hand-to-mouth activity reveals the number of events where hands (or objects) come into contact with the lips, but not necessarily resulting in entry into the mouth, then the Agency should consider separating their estimate of frequency of events into two terms. One term would represent the number of events where body parts and objects contact the lips and another term would represent the fraction of these events that result from contact with the interior of the mouth. The former is directly available from empirical data, whereas the latter may or may not be and consequently, one must rely on assumptions. Separation into two distinct terms may allow a less ambiguous use of empirical data and provide more transparency of that we know well and that we may not know well.

Unfortunately, there are few data on older infants and young toddlers. The Dutch study cited by the Agency has the only relevant data on these very young children (5 children, 3 to 6 months; 14 children, 6 to 12 months). It only reports duration of mouthing activities with no means of determining frequency of mouthing activities. No data are presented on activities that might contaminate the child's hand or what types of objects are put into the mouth. In addition, no data are presented on ethnicity, gender, region, season of year or other conditions indicating that these few children are representative of a whole population.

The use of 20 events per hour for hand to mouth events is a reasonable 90th percentile for older toddlers and preschool children based on the work of Zartarian and Reed. However, it is not protective since many of the children sampled had much higher rates (up to 62 hand-to-mouth and 39 objects-to-mouth in Reed). Furthermore, object-to- mouth events are not represented in this number, and with data from Gurunatnan (1998), it is clear that this can be a very important route of exposure. At present, there are few data available on 8 to 18 month-old children to determine if this is a meaningful frequency for the younger age group. Thus, a more realistic and protective cutoff of 95% should be used.

Also, one Panel member did not view the 20 events per hour as a reasonable 90th

percentile for older toddlers and preschool children stating that sample size in all ages in extremely small, variance is extremely large, the samples are too small to reflect differences (i.e., seasonal or regional behavior) and do not look at intra-individual variation. The absence of sufficient data, specifically for children 8 to 18 months was noted. One Panel member stated that since the sample size of the hand-to-mouth activity was small, the Standard Deviation associated with absorption would be large and that the upper range of the distribution should be used to add an additional precaution. Further, a 90th percentile cut off is not considered sufficiently conservative and protective for a "screening" function, because by definition it leaves 10% of the population unprotected.

The surface area noted for fingers would be appropriate for toddlers. The surface area of three fingers for a 15-kg child is not appropriate for all children. Additional information is needed about the portion and surface area of hand that the younger children put in the mouth. Very young infants can put a larger portion of the hand in the mouth than toddlers, but of course the hand is smaller. Whether 20 cm^2 is appropriate for the younger children needs to be determined. One can not do a linear back calculation from the toddler data. Infants also suck toes and arms. To some extent it appears that children can be classified as “mouthers” or “non-mouthers.” This is not to say that there are young children who don’t put fingers in the mouth, only that there are differences in the frequency and duration of these activities across children. Over the short term (1-6 days), it is believed that the 20 events per hour is an appropriate measure; however, for longer term exposures of young children, we are concerned about the issue of developmental changes that may influence the appropriateness of that value. Longitudinal studies are beginning in Texas and Arizona this fall that will eventually provide some answers. There may be other studies on-going that also are directed towards resolving this issue.

On page 68 of the overview document, a discussion of Wester (not Webster) et al. misstates the percentage of chlorine in Aroclor mixtures as the percent PCB, and then misinterprets that quantity as indicative of mass rather than type of chemical on skin.

Two Panel members commented on the importance of the dorsal surface of the hand to babies as they are somewhat immobile, and they use this portion of the hand while on a flat surface like a floor. Page 64 of the Agency's background document suggests that infants are at less risk when then are not yet mobile, but the opposite is also true. Pediatricians teach parents to place their children on the floor as a safe place when playpens and cribs are not available. An infant placed on a contaminated surface or near a crack or crevice treated area might end up with a higher exposure because of his/her inability to move. Babies often have both dorsal and ventral contact with surfaces including hands, arms, trunk, abdomen etc., depending upon the temperature and their clothing. This assumption is not safe and data are definitely required for the younger children for the micro-analysis. In addition, it was pointed out that a baby’s feet are accessible for mouth contact. There are peak periods of teething, and finger-to-mouth contact will vary by tooth type.

Issue #4 -Estimating Exposures from Secondary Sources

4. Given the relatively low magnitude of exposures from track-in, bathing or

showering relative to other scenarios, should OPP estimate exposure to pesticides that may result from these sources? If so, have we identified the most critical scenarios and approaches to be used to do the estimation?

Data are insufficient to state that exposures from bathing, showering or "track in" are of "relatively low magnitude" and should be minimized. Applications to homes, schools and yards do not result in a uniform distribution of residue levels, and the Agency should focus on post-application exposures from these types of scenarios. Thus, a "total daily exposure estimate" would be provided for assessing these types of exposures.

One Panel member thought that the primary focus of estimating exposures should focus on post application aspects of lawn or broadcast applications, providing the Agency with a "total daily exposure estimate".

Generally, the contact frequency for small children will be considerably greater in home than outdoors; it is not immediately obvious that the tracking-in exposure route will indeed be that inconsequential. There is insufficient information in the documents provided to determine whether track-in, bathing or showering pathways can be dropped a priori. One Panel member commented that residue levels resulting from "track in," showers or dust might become so minor that these potential routes of exposure would "drop out" (i.e., diminish to insignificant contributions to exposure). Pending the availability of better information that would justify ignoring these additional pathways, the Agency should continue working on developing models for estimating exposures by these routes.

Issue #5 - Estimating Exposures from Non-residential Pathways

5. OPP proposes to address exposure of children living on or near farms where pesticides are used by estimating deposition on lawns resulting from pesticide drift; OPP is developing a drift model for this purpose. Does the Panel consider this approach reasonable and are there other important non-residential pathways of potential pesticide exposure that should be evaluated for farm children?

Too few data are available to rely solely on pesticide residues on lawns in farming or adjacent residences to use in a drift model. A child's movement outdoors through pesticide-treated areas and deposition of residues indoors, from clothing worn by parents working in treated areas have to be considered. In addition, pesticide use patterns, housecleaning children's behavior in these environments must be addressed.

Concerns were raised that there are a variety of exposure pathways that represent potential exposures in addition to deposition on lawn. It was suggested that examples of children's exposures may include field exposures from walking to school through a field, from playing in fields along with working parents or on their own, from swimming in drainage ditches containing runoff, from residues entering a home, from residues on an applicator's clothing, and from residue and spills along paths and roadsides near houses or paths. While a previous presentation did not emphasize important exposures via drinking water, some areas are known to

be contaminated in farming communities and this should be considered.

It was pointed out that while the distribution of household dust and mass loading samples shown in Tables 31 and 32 of the Agency's background document are higher in the farmer and farm worker houses than in reference houses, the high range of samples were often more than a magnitude higher in these families. This speaks to the significant degree that some children might be at increased exposure. Also stressed was the importance of individual behaviors, including inappropriate use patterns, pica in children, and house cleaning. A study of lead-poisoned children was cited that found that none of the families with elevated blood lead levels had a vacuum cleaner in the house.

University of Washington researchers are currently investigating a population of agricultural community children. Residential proximity to active orchards is associated with higher body burdens (urinary organophosphate metabolites). Urinary metabolite levels in the children cannot be explained by contamination in diet, drinking water, house dust or indoor air. Exposure to spray drift (direct inhalation or dermal contact with deposited residues) is one possible explanation. Other hypotheses are also plausible.

(1) The child may be mobile and have access to treated areas, not just drift-impacted areas.

(2) Parental occupation in agriculture may result in “take home” of pesticides. (The actual mechanism involved is not well defined. Contaminated clothing is a plausible source of exposure to co-habitants of occupationally exposed persons.) Since conventional pathways do not predict observed biomonitoring results, additional pathways must be considered (and inclusion of spray drift in the SOP appears prudent). However, the research that should help clarify these exposure issues is ongoing. The missing pathway(s) has not been identified.

A number of general comments were made by the Panel members regarding estimating exposures from non-residential pathways. For example, one Panel member stated that there could be regional differences in exposures based upon commodities, application types, …. It was interjected that illegal pesticide applications could result in exposure. Mention was made of pesticide drift along the edge of fields and roadways that could result in exposures to children playing in these areas. One Panel member stated that pesticides had been found stored in well houses in North Carolina. As a result, North Carolina has data that demonstrates wells have sometimes become contaminated with pesticides. Another Panel member stated that children’s exposures in “hot spots” resulting from application drift might be more important than exposures from diet and dust. The Agency responded affirmatively to the question whether aerial, air blast, and boom sprayer applications were used in the exposure models.

Issue #6 - Addressing Exposure from Spray Drift

6. OPP is proposing to initiate the use of a spray drift model to estimate the likely magnitude of unintentional exposure to pesticide residues as a result of direct exposure to sprays. What is the Panel's opinion concerning the introduction of this new source of exposure into the risk assessment process?

Generally, the Panel concluded that using occupational exposures to determine exposure to non-occupational individuals was unrealistic. A question was raised as to how the model would deal with post-application exposure. Presently, the model does not have a means to include post-application exposures. In light of this, the use of dislodgeable residues might be used. The model assumes legal applications following label directions, but legal applications do not always happen and allowances should be made for this possibility in the model, (e.g., allow lawn residues to become dry before allowing children access).

Issue #7 - Twenty- four Hour Assumption Used in Estimating Risk in Schools, Day Care Centers, and Other Public Places

7. OPP currently assumes 24 hour residential exposure as a basis for its exposure assessments. OPP believes that this assumption is sufficiently conservative to protect from exposures that are likely to be encountered in other non-residential settings such as schools, day care centers, or other public places where the use patterns are comparable. Does the Panel agree or disagree and why?

Overall, since children's activities vary greatly between day care, home, school and other public places, the Agency should not assume a 24-hour residential exposure to be the sole basis for risk assessment. Micro-activity data, collected during school-time activities should be used as exposures might be greater in this environment compared to those in the home.

It is not a legitimate assumption that school, daycare and other public place exposures would be equal to or less than residential exposures. Activities of children vary greatly by location. School and daycare activities are substantially different from home activities. It seems that, for example, re-entry into a school after a three day weekend during which an insecticide was sprayed on Friday afternoon, might result in even higher exposures than re-entry into a home after 4-8 hours (Gurunathan, 1998). Furthermore, activities might well bring children into contact with residues at either higher or lower rates. Micro-activity data, if the model of micro-macro activity is found to be useful, would be necessary to decide on this issue. Since the number of items manipulated by children in school (e.g., computer keys, formica surfaces, vinyl tile floors) is greater than those in a house, it was felt that "micro" studies should be performed in schools.

One Panel member stated that, from an exposure standpoint, schools and homes were not equal. It was pointed out that individual variations existed regarding cleanliness in both schools and homes. It was mentioned that children at school could be playing on soils with exposures resulting from both dust and soil while they could be playing on grassy areas at home.

While use patterns might be comparable both in the physical environment and home, the behaviors and activities of the children are likely to be significantly different. It was noted that no supporting evidence was presented to back up the assumptions that activity patterns would be the same in both environments.

The use of videography in schools shows promise to obtain behavior patterns of interest that currently elude researchers. Videography in schools might actually be easier than in homes, because many schools are equipped with video equipment. A cross section of many children over several hours might enrich the database in terms of capturing intra- and inter-individual variation on hand-to-mouth, object-to-mouth and other behaviors of interest. Interesting things like the differences in hygiene, use of baby wipes, child- to-child contacts could be studied.

SAP Report No. 99-05B, November 18, 1999

REPORT:

FIFRA Scientific Advisory Panel Meeting, September 22, 1999, held at the Sheraton Crystal City Hotel, Arlington, Virginia

Session II - A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding:

Review of An Aggregate Exposure Assessment Tool

Mr. Paul I. Lewis
Designated Federal Official
FIFRA/Scientific Advisory Panel
Date:_____

Ronald J. Kendall, Ph.D.
Chair
FIFRA/Scientific Advisory Panel
Date:_____

Federal Insecticide, Fungicide, and Rodenticide Act
Scientific Advisory Panel Meeting
September 22, 1999

Session II: Review of An Aggregate Exposure Assessment Tool

PARTICIPANTS

Chair
Ronald J. Kendall, Ph.D., Professor and Director, The Institute of Environmental and Human Health, Texas Tech University/Texas Tech University Health Sciences Center, Lubbock, TX

FIFRA Scientific Advisory Panel
Charles C. Capen, D.V.M., Professor and Chairman, Department of Veterinary Biosciences The Ohio State University, Columbus, OH
Ernest E. McConnell, D.V.M., Toxpath, Inc., Raleigh, NC

FQPA Science Review Board Members
Scott Ferson, Ph.D., Applied Biomathematics, Setauket, NY
John C. Kissel, Ph.D., P.E., Associate Professor, University of Washington Department of Environmental Health, Seattle, WA
Ross B. Leidy, Ph.D., Director, Pesticide Residue Research Laboratory Department of Toxicology, North Carolina State University, Raleigh, NC
Nu-may Ruby Reed, Ph.D., Staff Toxicologist, California Environmental Protection Agency Department of Pesticide Regulation, Sacramento, CA
Lorenz Rhomberg, Ph.D., Gradient Corporation, Cambridge, MA

Designated Federal Official
Mr. Paul Lewis, FIFRA Scientific Advisory Panel, Office of Prevention, Pesticides and Toxic Substances, Environmental Protection Agency, Washington, DC

PUBLIC COMMENTS

Oral statements were received from:
Tim Pastoor, Ph.D. Robert Sielken, Ph.D. and David MacIntosh, Ph.D. (American Crop Protection Association)
David Burmaster, Ph.D. (Alceon Corporation)
Charles Benbrook, Ph.D. (Consumers Union)

Written statements were received from:
None were submitted

INTRODUCTION

The Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Scientific Advisory Panel (SAP) has completed its review of the set of scientific issues being considered by the Agency regarding the review of an aggregate exposure assessment tool. Advance public notice of the meeting was published in the Federal Register on September 3, 1999. The review was conducted in an open Panel meeting held in Arlington, VA, on September 22, 1999. The meeting was chaired by Ronald J. Kendall, Ph.D. Mr. Paul Lewis served as the Designated Federal Official.

One of the requirements of the 1996 Food Quality Protection Act specifies the Agency to consider aggregate exposure (i.e., exposure from residential, dietary and tap sources) to a pesticide. EPA, USDA and private organizations are participating in a co-operative agreement with Hampshire Research Institute (HRI) to develop a software package called LifeLine TM to perform aggregate (i.e. multiple sources and routes of exposure) exposure analysis.

Traditionally, exposure assessment has focused on characterizing the highest levels of exposure that will occur to an individual or a population over time as a result of the use of a pesticide. An alternative approach is to simulate the total dose received from multiple sources by individuals in a population through Monte Carlo analysis. The LifeLine TM model combines information on daily activity and dietary patterns from well known surveys to evaluate daily exposures to an individual. Transition rules (rules specifying how a value of an input is initially selected based on the characteristics of the population to be modeled and when and how the input values changes over time) will be applied. The Agency is seeking the Panel's advice on use of the software to calculate aggregate exposure.

Ms. Carol Christensen and Mr. Francis Suhre (Office of Pesticide Programs) discussed the goals and objectives of the Hampshire Research Institute Cooperative Agreement. Mr. Paul Price, John Young, Ph.D. and Christine Chaisson, Ph.D (Hampshire Research Institute LifeLine Project) provided a description of the LifeLine Project to model aggregate exposures to pesticides.

CHARGE

The specific issues to be addressed by the Panel are keyed to the background document , ***Review of an Aggregate Exposure Assessment Tool***, memorandum dated August 31, 1999 and are presented as follows:

1. The LifeLine^{TM} Model combines information on daily activity and dietary patterns from well known surveys to evaluate daily exposures to an individual and uses transition rules to specify how model input values are initially selected and when and how input values change over time.

1a. Does the Panel agree that the use of multiple databases (with appropriate transition rules) is a reasonable approach to modeling aggregate exposures?

1b. Would the Panel please comment whether aggregation of exposures needs to

consider demographic, spatial, and temporal factors affecting the exposure of an individual.

1c. Does the Panel view Microexposure Event Analysis with Monte Carlo techniques for incorporating probabilistic analysis as a viable foundation for a model of aggregate exposure?

2. Does the Panel have any practical suggestions on how to interpret and report results of aggregate daily exposure estimates to individuals over their lifetime, *e.g.*, to what extent should the exposure estimate be customized to fit the duration of the toxicological effect of each compound?

3. Would the Panel please comment on the key characteristics of the data sets currently used in the initial version of the LifeLine ™ Model?

4. Can the Panel suggest additional surveys and/or data sources for use in the LifeLine™ model (or other models of this type)?

PANEL RECOMMENDATION

1) Model Algorithm and Documentation: Maintain transparency and openness, with fully accessible source codes and clear transition rules. To the extent possible, make consistent the probabilistic analysis terminologies across Agency programs. There should be proper characterization of temporal correlations and associations, clear strategies for extending individual-based analysis to longitudinal assessment, and provision for uncertainty analysis that reflects current state of science. Obtain adequate validation and quality audit. Establish routine monitoring for software integrity and user input. Include scenario-specific case studies for future discussions.

2) Model Operation and Output: Include source contribution and sensitivity analysis. Design capability for modeling both the exposure and the absorbed dose, and for conducting analysis of different level of complexity (e.g., tier approach). Allow exposure estimates for various time periods as dictated by temporal, demographic, and spatial factors of exposure as well as toxicity characteristics that include pharmacokinetics and mechanisms of toxicity considerations.

3) Model Databank: Establish criteria for the choice of databases to be included in the model and ensure their mutual compatibility and capability to meet quality standards. Consider the use of a creditable consolidated database with clear pooling criteria. Add to the databank lists of Tolerances and Codex Maximum Residue Limits (MRLs), USDA PDP food residue database, and Agency data on pesticide use patterns that retain the resolution at state/regional level.

4) Additional Data Needs: For residential exposures, obtain data that can describe and predict dermal exposures from contact surfaces (e.g., transfer coefficients, dermal uptake, and contact duration) and data on patterns for non-agricultural uses. Consider data from Soil Contact Survey

I & II by Garlock et al. (J. Expos Anal Environ Epid, 9(2):134-142,1999). For drinking water exposure, obtain residue data in tap water.

DETAILED RESPONSE TO THE CHARGE

1. The LifeLine™ Model combines information on daily activity and dietary patterns from well known surveys to evaluate daily exposures to an individual and uses transition rules to specify how model input values are initially selected and when and how input values change over time.

1a. Does the Panel agree that the use of multiple databases (with appropriate transition rules) is a reasonable approach to modeling aggregate exposures?

The Panel acknowledged the effort to build a model for a more realistic analysis of aggregate exposures that will be made publicly available. The Panel agreed that no single database providing the adequate information exists. Multiple databases linked by transition rules was the only practical approach. The alternative, to make a grand scheme of random sampling of the entire U.S. population and all its activities, is not feasible. The Panel felt strongly that the results of the aggregate exposure must be related to a toxic endpoint (effect) for the analysis to have biological meaning.

It was pointed out that when using sizable databases from many sources, the criteria for the choice of database should be clearly stated. Care should be taken to ensure that the transition rules are valid and the databases are compatible (e.g., data fields and format). The model should be structured in a way to allow tracking of the contribution or influence of each database, especially those databases that are more lacking or limiting and when surrogate data are used.

For modeling aggregate exposures, it is very important that absorbed dose, not just the exposure, be modeled so that results can be compared to regulatory standards and biomonitoring data.

Credibility and adequate validation are important to a model. As the Panel also heard the testimony of the development of another program by the American Crop Protection Association (ACPA), the benefit of a synergistic collaboration between the Agency and ACPA should be considered. In general, there should be criteria for data quality, transparency in the model, and provision for quality auditing.

1b. Would the Panel please comment whether aggregation of exposures needs to consider demographic, spatial, and temporal factors affecting the exposure of an individual.

The software appears to be realistic in determining the total exposure to which one might be exposed based upon multiple routes and sources of exposure from published data. The emphasis on residential mobility is to be commended. Several questions were raised. Could the rapid changes in, for example, rural areas converting to subdivisions and databases (e.g., changes in pesticide classes) several years old have an effect on generated data? How will the software

handle the 1) differences between possible multiple exposures of inner-city inhabitants in New York vs. Chicago vs. Los Angeles and 2) differences in possible multiple exposures in urban middle-class neighborhoods vs. suburban middle-class neighborhoods? How will variations in interior environmental maintenance (i.e., cleaning patterns, indoor temperature maintenance, presence of pets and how they are cared for, traffic patterns) be addressed by the model? What are the sources of data on diet and activity patterns of infants and how are these records accessed? When looking at pesticide residues in food, how does the model address the potential for residues of compounds banned from use in commodities in the United States, versus their possible presence in imported products? How does one "customize" input to correspond with situations that do not fit the databases (e.g., a Caucasian married to a Latino who prepares, primarily, ethnic foods)?

The Panel agreed that demographic, spatial, and temporal factors and the correlation among behavior-dependent input variables should be addressed. In doing so, there is a tension between the degree of detail that one would like to divide the spectra into (e.g., fully accounting for variations among local regions, many narrow-range age groups) and the amount of data available to make separate accountings for the effects of these variables in different groups. It is important to remember that, to the extent that the groups identified retain internal heterogeneity, that heterogeneity becomes variation that the analysis has to acknowledge and deal with.

The individual-based analysis, insofar as modeling closer to the receptacles of exposure is more realistic than a combination of source-based models. The ultimate goal is to sufficiently characterize the exposure of individuals in a population and not to lose sight of the reasonable high-end individuals. Nevertheless, the individual-based analysis would still have to be linked back to the sources in terms of a clear identification of source contributions for risk presentation and the purpose of mitigation should it be necessary. The model would need to identify the source contribution at a certain benchmark percentile, e.g., the Agency's current policy of 99.9th percentile. Exposure data points at any specified percentile are not expected to all have the same main contributing sources.

1c. Does the Panel view Microexposure Event Analysis with Monte Carlo techniques for incorporating probabilistic analysis as a viable foundation for a model of aggregate exposure?

The Panel considers a longitudinal, individual-based approach using Monte Carlo simulation to be a viable foundation for the modeling and assessment of aggregate exposures. Indeed, it may well be the only practicable approach that is sufficiently general to make these complex temporal assessments. If chronic and long-term exposures are important considerations under FIFRA, as they no doubt should be, then such an approach should be central to a set of comprehensive assessment strategies. The Panel therefore meets the recent software development efforts in this direction with enthusiasm.

The most prominent advantage of this approach is its ability to make mechanistic estimates of cumulative aggregate exposures. The approach can explicitly account for long-term trends as well as seasonal and other cyclic fluctuations in the personal, dietary, and behavioral

characteristics of individuals being exposed. The approach could also include a variety of important temporal phenomena such as autocorrelation, elimination (excretion, depuration or detoxification), and perhaps even evolution of pesticide use patterns. Because the simulations create and check specific exposure scenarios, the results are likely to be extremely useful in tracking who is at risk of high exposure, and in suggesting potential ways that such risks might be reduced, with consequent improvement in the public health.

The Panel's enthusiasm is not without reservations, however. There are two substantial concerns about the approach, including 1) whether uncharacterized correlation patterns in real-world exposure scenarios strongly skew exposure distributions in a way that cannot be detected in assessments and 2) whether the absence of uncertainty propagation in the proposed approach renders the resulting exposure estimates unusable because they lack accompanying reliability statements. The Panel also makes several other comments relevant to the viability of the proposed approach as a foundation for aggregate exposure modeling.

Unrecognized associations and correlations

The world is a complex and highly multivariate place. Even the most relevant and extensive data sets are only caricatures of this complexity. The available data sets are grossly low-dimensional marginal projections of the actual patterns that exist. What we cannot see in such data sets are cryptic correlations and structure hidden in the high-dimensional patterns. Of course, the whole point of an individual-based longitudinal modeling approach is to incorporate as much of this structure as possible. Nevertheless, it is extremely difficult to completely capture all the important correlations and structure that are relevant. There are two mistakes that commonly arise in this circumstance leading to very different kinds of biases. The first mistake occurs when a characteristic of an individual that is actually varying through time is assumed to be constant over time. This mistake can result in grossly *over*estimating the dispersion of the final exposure. The other mistake occurs when a characteristic of an individual exhibits temporal autocorrelation yet is assumed to be fluctuating independently from one time step to the next. When this independence assumption (also called "Markovicity") is inappropriate, it can result in grossly *under*estimating the dispersion of the final exposure. These two mistakes represent poles on a spectrum where a correct model is somewhere in the middle. Many if not most individual characteristics relevant to exposure modeling are neither fixed through time nor independently fluctuating in each time step. The features in the LifeLine™ software that recognize several temporality categories (fixed, long-term trend, episodic, cyclic, and ephemeral) constitute a strategy to handle these intermediate cases when empirical information is available. It is not clear, however, what should be done when relevant empirical information is lacking. It may not be possible to bound the result simply by bracketing the assessment with a pair of simulations that either fix or independently vary the parameter. This issue requires further methodological research.

Uncertainty propagation

The modeling approach used in the LifeLine™ software is appropriate for handling *variability*, including temporal variability, in the dynamics of exposures. However, measurement error and related forms of *uncertainty* are ignored in the current approach. For instance, the plus-or-minus ranges which routinely accompany chemical determinations, are not incorporated into

the frequency distributions based on these determinations. In some cases, the magnitude of uncertainty may be very large, as in the self-reported bodily characteristics data. In other cases, a data set may have a higher precision with relatively little uncertainty. But some degree of incertitude is present in all measurements of course, and it is likely to have a bearing on the interpretation of the final assessment.

The approach should be augmented by methods that comprehensively propagate uncertainty as well as project variability. Attention should also be directed to how the reliability of the final results is best *expressed*. How should error bars be placed around the output summaries? Should lay and technical audiences be given different kinds of synopses? Because the assessments are, inescapably, extreme extrapolations from limited empirical information, it is crucially important to fully characterize the reliability of the final estimates. This is true a fortiori when the modeling addresses lifetime or long-term exposures.

It may turn out that honest propagation of the underlying measurement error and associated modeling uncertainties has a profound impact on some of the assessments developed with the proposed approach. It may, for instance, show that some assessments can make no precise determination about the likely exposures. This is not, in itself, a bad thing. We expect to be able to make only fairly poor predictions when empirical information is sparse and scientific understanding is limited. Indeed, the entire approach would be suspicious if it produced only apparent precision and surety in the face of massive underlying uncertainty. On the other hand, when exposures are either large enough or small enough relative to our uncertainties about them, the assessments will be reliable and can be used in regulatory determinations. In essence, a decision may not require further precision. It is not yet clear how many assessments will be reliable enough to be determinative in regulatory contexts and how many will be primarily useful for directing future empirical effort by highlighting data deficiencies. Which use can be made of an assessment's results can only be decided with appropriate uncertainty propagation.

Make use of established techniques

Individual-based modeling with Monte Carlo techniques has been used in biology for at least thirty years. Early examples were forest stand models in which individual trees were 'grown' inside a computer program to understand the dynamics of the stand. The early 1990s witnessed a blossoming of individual-based Monte Carlo models in used in ecological risk assessments. There is, as a result, a large literature that may be a useful resource for developers hoping to extend the individual-based approach to temporal exposure assessments. Indeed, the subject of Monte Carlo simulation has an extremely rich literature that offers a great many suggestions that might be used in the proposed approach. The Panel expects there may be many opportunities to exploit parallels between previous research and the current needs in exposure assessment methodology.

For example, the design of the LifeLine™ software appears to use a unitary structure in which a distribution of exposures is estimated under some assessment condition. Presumably, this distribution would then be compared with another baseline distribution without the contaminant (or pathway or whatever happens to be the focus of the particular assessment). It may be much more efficient instead to estimate the differential exposure in individuals. This can be

accomplished by simulating paired trajectories in which each individual has a twin, alike in all respects and experiencing the exactly same sequence of random fluctuations through time. The only difference between the two is that one assumes the presence of the contaminant (or pathway or whatever), and the other does not. With many such paired individuals followed through the simulations, it is easy then to compute the *distribution of differences* directly, rather than trying to compare two resulting distributions. This well known trick is called a "variance reduction" technique because it can significantly improve the precision of an assessment.

Software availability

It is essential that the methodological development be transparent and open. Electronic availability of the full source code, preferably over the Internet, will foster public trust in the approach and will also encourage parallel development by outside researchers that over the long run could add substantial value to the approach. These advantages are essential because formal beta review alone, however conscientious, is usually insufficient to uncover all the mistakes and obviate all the limitations in a newly developed approach to a complex problem. The requirement for openness and availability should apply to both the underlying software engine that manages the assessment and the assembled data sets and transition rules, insofar as they are codified. The Panel believes that such openness is not inconsistent with appropriate version control for the software.

Software integrity

Because validation in the strict scientific sense of checking conclusions and predictions against independent empirical data is generally impractical for a prognostic exposure assessment, the need to ensure integrity of the calculation tools is all the more important. Software development tools to check general software integrity should be routinely employed to confirm the absence of algorithm and coding faults such as range check errors, memory access violations, dimensional inconsistency, etc. Likewise, the interface design should include features to screen and echo all user input, and any units accompanying user input should be accepted and checked by the software.

The Panel is affirmative about the general principles but made some caveats about doing the analysis appropriately, such as properly accounting for correlations. As with any model, the Microexposure event approach is a simplification and abstraction of reality, and its value should not be judged on the degree to which it replicates reality, only on the degree to which it provides useful and usable insights into that reality. There are some aspects of the modeling that represent factors by mathematical representations that probably do not reflect the nature of the real processes, for instance, treating episodic factors as Markov processes (i.e., as having a constant per-time probability of changing). The real question is whether using this approach gives reasonable and useful values for the durations and prevalences of the factors in the real population. To the degree that these things can be checked, they would help validate the modeling approach. That is, one can validate the model not only on questions of pesticide exposure (e.g., by comparing to monitoring) but also on aspects of whether the distributions of exposure factors that the model predicts occur in practice.

It is not clear that the definition of Microexposure event analysis used in the presentation

is universally accepted. There may be some confusion on this point. Probabilistic approaches need not be truly micro-event oriented. Use of a probabilistic approach is clearly appropriate. Agency staff responsible for residential exposure assessment for pesticide registration distinguish macro (extended periods of relatively homogeneous behavior) and micro (truly event based) exposures. This distinction is useful. Some attempt at consistent use of terminology should be made.

Regarding the model operation, it was emphasized that the model with its capability for complex analysis should also have the flexibility to accommodate the various levels of exposure analysis, such as in a Tier Approach. There is an obvious advantage for using the same basic databases to conduct different tiers of analysis.

At this stage, the Panel discussion on the many aspects of the model is theoretical in nature. The expectation is that a future presentation of the model would include examples and case studies to demonstrate the use of the model. It will also be very useful to include comparisons of the model analyses to the estimation of aggregate exposure from the approach currently used by the Agency.

Question 2 :

Does the Panel have any practical suggestions on how to interpret and report results of aggregate daily exposure estimates to individuals over their lifetime, *e.g.*, to what extent should the exposure estimate be customized to fit the duration of the toxicological effect of each compound?

Lifetime doses are primarily of interest in cancer risk assessment. For many pesticides, cancer is not the toxicological endpoint of concern. Lifetime cumulative exposure prediction is impressive but of unknown utility at this point. Exposures over time frames pertinent to regulatory standards (acute, chronic within age classes) are of primary interest.

The Panel affirmed the importance of expressing the analysis of aggregate exposure based on both the toxicological information and the nature and pattern of exposures. Currently, the Agency exposure analysis generally addresses the dietary and drinking water exposure in terms of acute and chronic durations. On the other hand, residential exposures would likely to include short- and intermediate durations as well. Based on the use pattern for each pesticide, the presentation of the exposure analysis would have to include all these scenarios for all routes, with matching toxicity endpoints. The output should also include source contribution and sensitivity analysis.

Temporal issues are very important, and it is difficult to fully address this aspect without a clear idea of what is to be done in assessing the toxicity of temporally varying exposure, an aspect of the analysis that is not yet fully worked out. If the modeling is appropriately done, it will give simulated estimates of individual histories of time-varying exposure, and these histories can then be interpreted as to toxicological consequences, with averaging times made longer or shorter as is appropriate to the agent's pharmacokinetics and mechanisms of toxicity. The only issue about customizing, then, is to ensure that toxicities of interest can all be addressed with exposure data

on an appropriately fine scale. The main structure of the exposure modeling is to estimate daily exposure, but there is no reason in principle that finer time divisions could not be made for some kinds of exposure in cases where hourly or even minute-by-minute changes in exposure are toxicologically important.

Although toxic effects occur on different time scales, they all must relate back to the actual temporally varying patterns of exposure as experienced. A series of acute episodes may add up to a sub-chronic exposure if suffered in just the right sequence, and all exposures—short- and long-term—affect the chronic exposure. The averaging time that is appropriate may vary with toxicity, but the exposures should be characterized in the full richness of their temporal patterns to the extent they can be, leaving the challenge of interpreting the toxicological consequences of such patterns to the next step of the analysis. To the degree that toxicological characterization does not deal effectively with time-varying exposure, it may be hard to make appropriate interpretations of the time variation, but the overall rationale and structure of the methods should allow for improvement rather than making different exposure estimates for different toxicities.

Question 3 :

Would the Panel please comment on the key characteristics of the data sets currently used in the initial version of the LifeLine™ Model?

The Aggregate Exposure Tool as described is applicable to residential, dietary, and tap water exposures. Material presented is heavily oriented toward dietary exposures.

Existing data are inadequate to describe/predict exposures from dermal contact with residues on surfaces. Transfer coefficients are largely unknown. Dermal estimates are currently little more than guesswork. Data for characterizing contact time is also lacking.

Panel comments on the database are mainly focusing on the generic survey databases as listed in the presentation document. These included natality data, mobility, human activities, physiological parameters, and dietary patterns. In general, current data for dietary and drinking water consumptions are more abundant than for residential exposures. One issue is the compatibility of the survey data included in the model. For example, dietary exposures may be grouped into four U.S. regions while the census data would include a fifth region, "Pacific". Thus, adjustment would be necessary if analysis is grouped by the U.S. regions.

The various data sets are collected for different program purposes. Many other databases, some perhaps more limited in size and geographic locations, also exist. Compatibility should be considered in the criteria for database selection. In this regard, there may be advantages in using creditable consolidated databases with clear criteria for pooling data. One such example is the potential usefulness in Consolidated Human Activity Database (CHAD) that was developed to support the exposure assessment for the Office of Research and Development.

Question 4:

Can the Panel suggest additional surveys and/or data sources for use in the

LifeLine™ model (or other models of this type)?

Since the model is designed to be available for public use, it is desirable to include not only the generic databases as outlined in the model documentation (e.g., data for age-linked physiological parameters, dietary and drinking water consumptions, residential mobility and activity patterns) but also pesticide-specific databases. For dietary exposures, this would include data that have regulatory meanings, such as Tolerances and Codex Maximum Residue Limits. It was understood that residue data from USDA Pesticide Data Program will be a part of the model databank as well.

The Panel concluded that for Aggregate Exposure Assessment to be of value it has to be "believable" for it to be credible! The best way to establish credibility is to validate the assessment with actual data. This could be accomplished by choosing a chemical (not necessarily a pesticide) for which a biomarker of exposure is known and then "test" the exposure assessment. One would accomplish this by first conducting the aggregate exposure assessment and then comparing the results to the "marker" of exposure to see how close the fit actually is. If the fit is close, then one can have confidence in the assessment. Conversely, if it is not close, then one would have to couch the results of the assessment with some type of confidence limit. Another result of such an analysis would be that the validation process would identify the weaknesses in the assessment so that they could be addressed in future assessments. This iterative process would then ideally evolve into a credible aggregate exposure assessment.

Data for pesticide use patterns compiled by the Agency should be also included. These data are crucial in refining analyses for all pathways of exposure, and yet the public generally do not have access to these types of data. Since pesticide use pattern could vary widely from location to location, it is desirable not to compact the data to an extent of losing the characteristics of the pattern for a specific location or region. While exposure analysis that provides a nationwide picture is important, the state and/or regional patterns should also be retained. Collection of data for the pattern of non-agricultural uses is needed since these data are generally more lacking than data for agricultural uses. Since these exposures tend to be episodically high, they would be very useful for a realistic estimation of acute exposures and risk.

As a rule, residue data closer to what people are coming in contact with or exposed to are more desirable. In this regard, one major area of data needs is the database for drinking water. The current models for estimating pesticide concentrations in water are useful for screening analyses. The well and surface water surveys have also been useful. However, survey data on tap water would most closely reflect the concentrations in the water people drink. The effort to obtain tap water data is particularly encouraged. One possibility for obtaining this type of data may be through collaboration with the routine tap water surveys conducted by many municipal drinking water suppliers.

Regarding residential exposures, one other database that may potentially be useful for the aggregate exposure analysis is the Soil Contact Survey I & II by Garlock et al. (J. Expos Anal Environ Epid, 9(2):134-142,1999).

At the conclusion of the Panel discussions, it was restated that the review of LifeLine™ program by the Panel does not constitute an endorsement of the program itself.

LITERATURE CITED

Garlock et al. *Soil Contact Survey I & II*. J. of Expos Anal Environ Epid, 9(2):134-142,1999).

.

SAP Report No. 99-05C, November 18, 1999

REPORT:

FIFRA Scientific Advisory Panel Meeting, September 22, 1999, held at the Sheraton Crystal City Hotel, Arlington, Virginia

Session III - A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding:

Identifying Carbamate Pesticides that Have a Common Mechanism of Toxicity

Mr. Paul I. Lewis
Designated Federal Official
FIFRA/Scientific Advisory Panel
Date:_____

Ronald J. Kendall, Ph.D.
Chair
FIFRA/Scientific Advisory Panel
Date:_____

Federal Insecticide, Fungicide, and Rodenticide Act
Scientific Advisory Panel Meeting
September 22, 1999

Session III: Identifying Carbamate Pesticides that Have a Common Mechanism of Toxicity

PARTICIPANTS

Chair
Ronald J. Kendall, Ph.D., Professor and Director, The Institute of Environmental and Human Health, Texas Tech University/Texas Tech University Health Sciences Center, Lubbock, TX

FIFRA Scientific Advisory Panel
Charles C. Capen, D.V.M., Professor and Chairman, Department of Veterinary Biosciences The Ohio State University, Columbus, OH
Ernest E. McConnell, D.V.M., Toxpath, Inc., Raleigh, NC

FQPA Science Review Board Members
William Brimijoin, Ph.D., Director, Department of Pharmacology, Mayo Clinic and Medical School, Rochester, MN
Gary P. Carlson, Ph.D., Professor of Toxicology and Associate School Head, School of Health Sciences, Purdue University, West Lafayette, Indiana
Rory Conolly, Ph.D., Senior Scientist, Chemical Industry Institute of Toxicology, Research Triangle Park, NC
David Dorman, Ph.D., DVM, Chemical Industry Institute of Toxicology, Research Triangle Park, NC
Amira Eldefrawi, Ph.D., Professor, Department of Pharmacology & Experimental Therapeutics, University of Maryland School of Medicine, Baltimore, MD
Carey Pope, Ph.D., School of Pharmacy, Northeast Louisiana University, Monroe, LA
Lorenz Rhomberg, Ph.D., Gradient Corporation, Cambridge, MA
Stephen M. Roberts, Ph.D., Professor and Program Director, University of Florida Center for Environmental & Human Toxicology, Gainesville, FL
Lester Sultatos, Ph.D., Department of Pharmacology and Physiology, New Jersey Medical School, Newark, NJ
Kendall B. Wallace, Ph.D., Scientific Advisor, University of Minnesota, Duluth, MN

Designated Federal Official
Mr. Paul Lewis, FIFRA Scientific Advisory Panel, Office of Prevention, Pesticides and Toxic Substances, Environmental Protection Agency, Washington, DC

PUBLIC COMMENTS

Oral statements were received from:
Karin Bentley (American Crop Protection Association)

Written statements were received from:
None were submitted

INTRODUCTION

The Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Scientific Advisory Panel (SAP) has completed its review of the set of scientific issues being considered by the Agency regarding the review of an aggregate exposure assessment tool. Advance public notice of the meeting was published in the Federal Register on September 3, 1999. The review was conducted in an open Panel meeting held in Arlington, VA, on September 22, 1999. The meeting was chaired by Ronald J. Kendall, Ph.D. Mr. Paul Lewis served as the Designated Federal Official.

The 1996 Food Quality Protection Act requires the Agency to consider the cumulative effects of exposure to pesticides sharing a common mechanism of toxicity. The Office of Pesticide Programs (OPP) is developing guidance for conducting cumulative risk assessment, this guidance will be presented later at this SAP meeting (see SAP Report No.99-05D). The identification of a candidate group of pesticides and other substances that cause a common mechanism is the first step in the cumulative risk assessment process. Since organophosphorus and carbamate pesticides have been assigned priority for tolerance reassessment, OPP has considered data for identifying whether these pesticides cause common effects by common mechanism of toxicity. Inhibition of acetylcholinesterase (AChE) has been a focal point for common toxic effects of most organophosphorous and many carbamate pesticides.
William Sette, Ph.D. served as the lead Agency presenter soliciting the Panel's advice concerning OPP's conclusion that AChE inhibition provides a sufficient basis for determining a common mechanism of toxicity for grouping carbamate pesticides. In addition, the Agency requested the Panel's comment on whether carbamate pesticides that inhibit AChE should be grouped with the organophosphate pesticides that also cause AChE inhibition.

CHARGE

The specific issues to be addressed by the Panel are keyed to the background document, ***Identifying Carbamate Pesticides that Have a Common Mechanism of Toxicity***, memorandum dated August 31, 1999 and are presented as follows:

1. Does the Panel agree with the Agency's conclusion that acetylcholinesterase inhibition provides a sufficient basis for determining a common mechanism of toxicity for grouping carbamate pesticides?

2. Does the Panel agree with the Agency's conclusion that carbamate and OP pesticides which inhibit AChE should be considered to be a single common mechanisms grouping?

PANEL RECOMMENDATION

Although the capacity to inhibit acetylcholinesterase is an appropriate mechanism for the initial grouping of carbamates and organophosphorous pesticides, other critical factors such as potency, rates of reactivation of inhibited enzyme, recovery rates *in vivo*, and non-cholinergic effects of these compounds should be considered for the final grouping of pesticides that will undergo a cumulative risk assessment.

Efforts should be made to acquire data on the effects of combinations of carbamate and organophosphorous pesticides at individual levels of exposure of each pesticide in the range of their NOELs.

DETAILED RESPONSE TO THE CHARGE

1. Does the Panel agree with the Agency's conclusion that acetylcholinesterase inhibition provides a sufficient basis for determining a common mechanism of toxicity for grouping carbamate pesticides?

It is assumed that the term "carbamate" refers also to the thio- and di-thiocarbamates. The Panel agreed unanimously with the Agency's conclusion that acetylcholinesterase inhibition provides a sufficient basis for determining a common mechanism of toxicity for grouping carbamate pesticides. Those carbamates that inhibit acetylcholinesterase can lead to an accumulation of acetylcholinesterase at muscarinic and nictotinic receptors, thereby leading to symptoms of cholinergic crisis. Excessive activation of muscarinic receptors can produce diarrhea, abdominal cramps, miosis, excessive secretions of endocrine and exocrine glands, and lowered heart rate and blood pressure. In contrast, excessive acetylcholine at nicotinic receptors will produce receptor desensitization, leading to muscle paralysis. While all anticholinesterase carbamates can, at least in theory, lead to these effects, many Panel members emphasized that some carbamates are only weak inhibitors of acetylcholinesterase and that potency and dose will be extremely important factors when considering which carbamates might be grouped for a final cumulative risk assessment.

Additionally, a Panel member noted the importance of clarifying the criteria for establishing whether or not a compound inhibits acetylcholinesterase. For example: Must *in vivo* or *in vitro* inhibition be established? What enzyme sources are acceptable with regard to species and tissues? Is inhibition of butyrylcholinesterase equivalent to inhibition of acetylcholinesterase? Should solubilization of the enzyme be required? What are the specific assay conditions?

The Panel concluded that there was a strong need to validate the potential additivity of exposure to carbamates. While there are some data with regard to combinations, most of it is with "high" exposures. Because most exposures of concern will be low, a well-controlled experiment at low exposure levels of carbamates in animals is warranted and strongly encouraged. A similar study with combinations of carbamates and organophosphorous pesticides would be most useful to establish interactions of these chemicals.

The Panel also agreed that groupings of carbamates based on non-cholinergic endpoints such as reproductive, thyroid, developmental, and broad-spectrum neurotoxicity could possibly be appropriate for certain carbamates, especially the low-potency, thio- and di-thiocarbamate fungicides and herbicides, whose ability to inhibit acetylcholinesterase is weak or absent.

2. Does the Panel agree with the Agency's conclusion that carbamate and OP pesticides which inhibit AChE should be considered to be a single common mechanisms grouping?

The Panel agreed unanimously with OPP's conclusion that carbamate pesticides that inhibit acetylcholinesterase should be grouped with the organophosphorous pesticides that also cause acetylcholinesterase inhibition. However, the Panel also agreed that differences in potency (particularly among the carbamates) and in rates of reactivation of carbamylated versus phosphorylated acetylcholinesterase, as well as differences in time to recovery *in vivo*, are critical factors that should play a role in conducting the actual estimate of cumulative risk. The Panel also agreed that risk assessments now have to focus on endpoints (and all the agents that cause that endpoint) rather than on agents (and all the endpoints caused by the agents). It would be a mistake to apply the cumulative risk umbrella to chemical classes of agents per se; rather one should apply it to all agents that act through the common mechanism.

One Panel member pointed out that a recent International Life Sciences Institute (ILSI) report by Mileson *et al* (1998) lists three criteria for determining a common mechanism of toxicity: Do the agents cause the same critical effects? Do they act on the same molecular target at the same target tissue? Do they act by the same biochemical mechanism of action (perhaps by sharing a common toxic intermediate)? While all three criteria are met by the carbamate pesticides as a group and by the organophosphorous pesticides as a group, an argument can be made that the third criteria is not satisfied when these two classes of agents are considered together. It was noted, however, that the ILSI report did not reach a consensus on whether compounds could share a common mechanism of toxicity if only two of the three criteria are met.

Panel members also stated the importance of empirical studies in animal models directed toward characterizing acetylcholinesterase inhibition following exposures to combinations of carbamates and organophosphorous pesticides, with individual pesticide exposures in the range of each NOAEL. Currently such data are lacking.

SAP Report No. 99-05D, November 18, 1999

REPORT:
FIFRA Scientific Advisory Panel Meeting, September 23, 1999, held at the Sheraton Crystal City Hotel, Arlington, Virginia

Session IV- A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding:

Proposed Guidance for Conducting Cumulative Hazard Assessments for Pesticides that Have a Common Mechanism of Toxicity

Ms. Laura E. Morris
Designated Federal Official
FIFRA/Scientific Advisory Panel
Date:_____

Ronald J. Kendall, Ph.D
Chair
FIFRA/Scientific Advisory Panel
Date:_____

Federal Insecticide, Fungicide, and Rodenticide Act
Scientific Advisory Panel Meeting
September 23, 1999

SESSION IV - Proposed Guidance for Conducting Cumulative Hazard Assessments for Pesticides that Have a Common Mechanism of Toxicity

PARTICIPANTS

Chair
Ronald J. Kendall, Ph.D, Professor and Director, The Institute of Environmental and Human Health, Texas Tech University/Texas Tech University Health Sciences Center, Lubbock, TX

FIFRA Scientific Advisory Panel
Ernest McConnell, DVM, Toxpath, Inc., Raleigh, NC

FQPA Science Review Board Members
Sharon Black, D.V.M., Ph.D., DACVP, Associate Professor, College of Veterinary Medicine, Mississippi State University, Starkville, MS
Rory Conolly, Ph.D., Senior Scientist, Chemical Industry Institute of Toxicology, Research Triangle Park, NC
Lorenz Rhomberg, Ph.D., Gradient Corporation, Cambridge, MA
Stephen Roberts, Ph.D., Professor and Program Director, University of Florida Center for Environmental & Human Toxicology, Gainesville, FL
Jerrold Ward, D.V.M., Ph.D., Chief, Veterinary and Tumor Pathology Section, National Cancer Institute, Frederick, MD
John Cullen, V.M.D., Ph.D., Professor of Pathology, College of Veterinary Medicine, North Carolina State University, Raleigh, NC

Designated Federal Official
Ms. Laura Morris, FIFRA Scientific Advisory Panel, Office of Prevention, Pesticides and Toxic Substances, Environmental Protection Agency, Washington, DC

PUBLIC COMMENTERS

Oral statements were made by:

Charles M. Benbrook, Consumers Union

Daniel M. Byrd III, Ph.D., D.A.B.T., Consultants in Toxicology, Risk Assessment and Product Safety

Resha Putzrath, Principal for Toxicology, Risk Assessment and Risk Communication, Georgetown Risk Group and Johns Hopkins University

Written statements were received from:

Daniel M. Byrd III, Ph.D., D.A.B.T., Consultants in Toxicology, Risk Assessment and Product Safety

Greg R. Christoph, Ph.D., E.I. du Pont de Nemours and Company, Central Research and Development, Haskell Laboratory, Newark, DE

INTRODUCTION

The Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Scientific Advisory Panel (SAP) has completed its review of the set of scientific issues being considered by the Agency regarding issues pertaining to the assessment of residential exposure to pesticides. Advance notice of the meeting was published in the *Federal Register* on September 3, 1999. The review was conducted in an open Panel meeting held in Arlington, Virginia, on September 21, 1999. The meeting was chaired by Ronald J. Kendall, Ph.D, Professor and Director, The Institute of Environmental and Human Health, Texas Tech University/Texas Tech University Health Sciences Center, Lubbock, TX. Ms. Laura Morris served as the Designated Federal Official.

The 1996 Food Quality Protection Act requires the Agency to consider the effects of cumulative exposure to pesticides sharing a common mechanism of toxicity. Cumulative risk assessment encompasses the integration of the hazard potential of non-occupational exposure in the aggregate (i.e., multi-pathway and route) and focus only on those pesticides which share a common mechanism of toxicity. For this session, the Agency will provide an overview of the hazard and dose response component and a case study. The exposure and risk characterization component will be presented at the December 8-9, 1999 SAP meeting. When completed, the Office of Pesticide Programs (OPP) will use the guidance document for conducting cumulative risk assessments for those chemicals that are toxic by a common mechanism.

The Agency seeks SAP input on this guidance document in the following areas: (1) selection of chemicals for a cumulative risk assessment; (2) selection, normalization, and adjustment of the point of departure for cumulating the common toxicity; (3) incorporation of group uncertainty factors; (4) methods for estimating the cumulative toxicity and; (5) review of OPP's case study.

CHARGE

The specific issues to be addressed by the Panel are keyed to the background document, ***Proposed Guidance on Cumulative Risk Assessment of Pesticide Chemicals: Issues Pertaining to Hazard and Dose-Response***, memorandum dated September 22, 1999, and are presented as follows:

Issue 1. Selection of Chemicals for a Cumulative Risk Assessment

Chapter 3 of the background document emphasizes that all chemicals which have been initially grouped by a common mechanism of toxicity are not necessarily appropriate for inclusion in a final cumulative risk assessment. There are both hazard and exposure considerations.

1. Does *Chapter 3* clearly present additional hazard considerations that are needed to determine those chemical members which should be included in the final cumulative risk assessment?

Issue 2. Selection, Normalization, and Adjustment of the Point of Departure (POD) for Cumulating the Common Toxicity

As discussed in *Chapter 5.1-5.2* of the background document, a point of departure (i.e., a dose or exposure metric corresponding to some fixed marker of toxicity) must be selected to sum the combined exposure for the chemical group. To the extent possible, the PoDs should reflect a uniform measure of the common toxic effect, which is produced by a common mechanism of toxicity, across the chemical members. A benchmark dose approach is preferred to derive the PoDs for each chemical member.

In single chemical assessments, the Agency uses the upper bound estimates (i.e., the lower confidence limit on dose) for both cancer (called LED) and noncancer benchmark dose assessment. The concern has been raised, however, that summing upper bounds of multiple compounds may result in a exaggerated risk.

2. Does the SAP agree that it is more appropriate to sum the central estimates (i.e., ED) rather than combining upper bounds in the cumulative risk assessment of multiple chemicals?

Issue 3. Incorporation of Group Uncertainty Factors

As discussed in *Chapter 5.3* of the background document, traditionally one or more of the uncertainty factors are used to derive a Reference Dose (RfD) for a single chemical. There are five uncertainty factors that are considered to account for the following extrapolations: LOAEL to NOAEL (UF_L), subchronic NOAEL to chronic NOAEL (UF_S), experimental animal to humans (UF_A), interhuman variation (UF_H), and incomplete database to complete database (UF_D). It is proposed that the extrapolations of LOAELs to NOAELs or subchronic NOAELs to chronic NOAELs be applied as adjustments of a chemical's PoD before estimating the cumulative risk. These adjustments are meant to be based on some scientific data that permit a reasonable extrapolation or interpolation rather than applied solely as a science policy default decision. It is further proposed that other traditional uncertainty factors be treated as a composite "group uncertainty factor" that pertains to the chemical members as a whole. Thus, the intra- and inter-species UFs and the database completeness UF are applied as a composite group factor after cumulative risk is estimated (i.e., not before on each chemical's PoD). The rationale of the group UF is based on the premise that these factors should be viewed for the group as a whole, given that all the chemicals are anchored by a common toxic effect produced by a common mechanism. Additionally, one is not simply evaluating risk in the context of a single chemical data base but the database for all the chemicals in the assessment. The advantage of a group uncertainty factor is that it allows one to separate the resulting risk that is based on scientific adjustments from judgmental policy decisions to account for uncertainty. It is also further proposed that a FQPA Safety Factor decision be applied for the group rather than on individual pesticides.

3. Does the SAP agree with this approach, and does the background document clearly describe the rationale and guidance for the implementation of chemical specific adjustment factors and of a group UF for the cumulative risk assessment? Has the document clearly presented the

limitations and strengths of the group UF approach?

Issue 4. Methods for Estimating the Cumulative Toxicity

As discussed in *Chapter 5.6* of the background document, one of the steps in the cumulative risk assessment process will be to select a method to cumulate dose or exposures. This method will serve to normalize differences in the toxic potencies among the chemicals in the cumulative assessment. Precedence in the Agency's 1986 and revised 1999 *Guidance for Conducting Health Risk Assessment of Chemical Mixtures* describes several techniques for estimating risk to multiple chemicals. The cumulative guidance focuses on the component-based dose addition methods used in the EPA's chemical mixture assessment guidance document. Two methods, a margin-of-exposure approach and an approach using relative potency factors, are presented.

4a. Does the SAP agree that both methods are valid to consider for estimating cumulative risk associated with exposures to chemical that cause a common toxic effect by a common mechanism? Has the document clearly described these two approaches and their strengths and limitations? Are there other methods that OPP should consider?

4b. It is anticipated that most mechanisms of toxicity encountered currently will be nonlinear dose-response relationships. Nevertheless, for mechanisms of toxicity consistent with linear dose-response relationships, does the SAP agree that using the relative potency factor approach by summing the slopes of the dose-response curves is an appropriate method? If not, what methods would the SAP recommend for low-dose linear extrapolations of risk?

Issue 5. Case Study

In *Appendix A* of the guidance document is a case study on organophosphorous pesticides.

5. Does this case study provide a clear example of the application of the hazard and dose-response elements of the draft guidance?

PANEL RECOMMENDATION

Validation of the assessment methodology could be accomplished by choosing a chemical for which a biomarker of exposure is known and then "test" the exposure assessment. One would first conduct the cumulative exposure assessment and then compare the results to the "marker" of exposure to see how close the fit actually is. If the assessment is close to the observed data, then one can have confidence in the assessment. Conversely, if it is not close, then one would have to express the results of the assessment with some type of confidence limit. Another result of such an analysis would be that the validation process would identify the

weaknesses in the assessment methodology so that they could be addressed in future assessments. This iterative process would then evolve into a credible exposure assessment.

DETAILED RESPONSE TO THE CHARGE

Issue 1. Selection of Chemicals for a Cumulative Risk Assessment

Chapter 3 **of the background document emphasizes that all chemicals which have been initially grouped by a common mechanism of toxicity are not necessarily appropriate for inclusion in a final cumulative risk assessment. There are both hazard and exposure considerations.**

1. Does *Chapter 3* clearly present additional hazard considerations that are needed to determine those chemical members which should be included in the final cumulative risk assessment?

Chapter 3 of the background document, Hazard Assessment and Characterization, adequately presents hazard considerations for inclusion of a chemical in a Common Mechanism Group (CMG). While many of the additional hazard considerations for inclusion of a chemical in a Cumulative Assessment Group (CAG) are also presented, members of the SAP recognize that there are other additional hazard considerations remaining to be clearly defined in order for chemicals to be appropriately selected for a CAG. Factors to consider in the refinement of additional hazard considerations include the following:

a. There should be consideration of issues of exposure and the likelihood of co-exposure to chemicals in a CMG.

b. There should be consideration of the pharmacokinetics (PK) and pharmacodynamics (PD) of the entire mechanism of action of each chemical in a CMG in order to distinguish between common PK and common PD when making comparisons within the CMG.

c. It should be determined if dose-response data for each chemical are adequate to allow for 1) an acceptable degree of confidence in points-of-departure and 2) assessment of whether or not individual chemicals have parallel dose-response curves. Parallel dose-response curves are required for dose-addition, the default method for estimation of cumulative risk in the proposed guidance.

Additional specific recommendations from the Panel are as follows:

a. In order to be consistent with terminology used in the FQPA, the Agency erroneously uses "mechanism of toxicity" and "mode of action" as equivalent terms in Chapter 3 instead of using the terms according to their commonly accepted definitions. It is recommended that the Agency use "mode" to address what FQPA calls "mechanism," and the Guidance should justify

the Agency's usage of conventional terminology.

b. The selection process used to identify chemicals for inclusion in a CMG should be subject to external peer review (given the possible significant regulatory consequences of selection to a CMG).

A Panel member commented that at this stage the Agency cannot be certain that all potentially relevant considerations have been identified, including those considered to be "the simplest case" (i.e., acetylcholinesterase-inhibiting pesticides). For example, one crucial question is whether acetylcholinesterase-inhibition or unrelated action is the dominant mode of toxicity. Secondly, what must the relative potencies of two such actions be for the compound in order to qualify for inclusion in the CAG. General guidelines from the Agency regarding this point would be beneficial.

Issue 2. Selection, Normalization, and Adjustment of the Point of Departure (POD) for Cumulating the Common Toxicity

As discussed in *Chapter 5.1-5.2* of the background document, a point of departure (i.e., a dose or exposure metric corresponding to some fixed marker of toxicity) must be selected to sum the combined exposure for the chemical group. To the extent possible, the PoDs should reflect a uniform measure of the common toxic effect, which is produced by a common mechanism of toxicity, across the chemical members. A benchmark dose approach is preferred to derive the PoDs for each chemical member.

In single chemical assessments, the Agency uses the upper bound estimates (i.e., the lower confidence limit on dose) for both cancer (called LED) and noncancer benchmark dose assessment. The concern has been raised, however, that summing upper bounds of multiple compounds may result in a exaggerated risk.

2. Does the SAP agree that it is more appropriate to sum the central estimates (i.e., ED) rather than combining upper bounds in the cumulative risk assessment of multiple chemicals?

The Panel agreed that, in order to avoid compounding conservativism, it is appropriate to sum the central estimates (EDs) rather than combining upper bounds in the cumulative risk assessment of multiple chemicals. For the particular purpose of assessing relative potency, central estimates are better since they provide unbiased estimates of the relative contribution of components of a mixed exposure. The issue is not just conservatism with regard to overall risk, but is also determination of the contributions of each agent in the right proportions. Depending on the exposure levels, it is not necessarily conservative to use LEDs for each agent when doing relative potency as one can overemphasize the contribution of a low-potency, low-exposure chemical and hence under emphasize the risk from a more potent, higher exposure one.

Further, a Panel member noted that the proposed guidance did not clearly establish

whether the effect of using the lower confidence limit for each agent would significantly increase the safety factor, and if so, by how many fold. The derivation of a reference dose (using ED as a benchmark for cumulative risk from two agents) higher than the most potent of two compounds used in combination (moderately high potency and moderately low potency) would be perverse. More detailed consideration of this issue needs to be included in the guidance document.

This question raised the following concerns from Panel members:

a. Using central estimates does not provide a means for considering uncertainty associated with variability in the dose-response data; the proposed guidance intends to address this through application of a database uncertainty factor (UFD) after assessment of cumulative hazard; will the method of selecting this UFD be scientifically sound enough to make use of central estimates defensible?

b. The proposed guidance does not thoroughly address the impact of using upper bound estimates in cumulative risk assessment; i.e., to what degree would the safety factor potentially be exaggerated?

Issue 3. Incorporation of Group Uncertainty Factors

As discussed in *Chapter 5.3* of the background document, traditionally one or more of the uncertainty factors are used to derive a Reference Dose (RfD) for a single chemical. There are five uncertainty factors that are considered to account for the following extrapolations: LOAEL to NOAEL (UF_L), subchronic NOAEL to chronic NOAEL (UF_S), experimental animal to humans (UF_A), interhuman variation (UF_H), and incomplete database to complete database (UF_D). It is proposed that the extrapolations of LOAELs to NOAELs or subchronic NOAELs to chronic NOAELs be applied as adjustments of a chemical's PoD before estimating the cumulative risk. These adjustments are meant to be based on some scientific data that permit a reasonable extrapolation or interpolation rather than applied solely as a science policy default decision. It is further proposed that other traditional uncertainty factors be treated as a composite "group uncertainty factor" that pertains to the chemical members as a whole. Thus, the intra- and inter-species UFs and the database completeness UF are applied as a composite group factor after cumulative risk is estimated (i.e., not before on each chemical's PoD). The rationale of the group UF is based on the premise that these factors should be viewed for the group as a whole, given that all the chemicals are anchored by a common toxic effect produced by a common mechanism. Additionally, one is not simply evaluating risk in the context of a single chemical data base but the database for all the chemicals in the assessment. The advantage of a group uncertainty factor is that it allows one to separate the resulting risk that is based on scientific adjustments from judgmental policy decisions to account for uncertainty. It is also further proposed that an FQPA Safety Factor decision be applied for the group rather than on individual pesticides.

3. Does the SAP agree with this approach, and does the background document

clearly describe the rationale and guidance for the implementation of chemical specific adjustment factors and of a group UF for the cumulative risk assessment? Has the document clearly presented the limitations and strengths of the group UF approach?

The Panel generally supports the use of group uncertainty factor (UF) approach as presented in the proposed guidance. The Panel recognizes the effort made in the Guidance Document to distinguish aspects of the UFs that adjust results to put them on a comparable footing (and thus should go before combination) from those aspects that express uncertainty about application of the final result to human risk (and hence should go after combination). This distinction represents an important advancement by the Agency and a contribution to the development of cumulative risk methodology. It would be helpful if a further discussion pointed out that all uncertainty factors are partially a means of extrapolation (i.e., applying a needed adjustment) and partially an allowance for the uncertainty in the size of that adjustment. The method proposed in the document emphasizes the extrapolation aspect of factors for LOAEL-to-NOAEL and subchronic-to-chronic factors in that they are needed to adjust different studies to put the endpoints on equal footing before combination. It emphasizes the allowance-for-uncertainty aspect of the inter- and intra-species adjustments and completeness factor in that these are a product of the uncertainty in relative effect of the agent among species and among differently sensitive humans. In fact, both sets of factors may contain some aspects of adjustment and some aspects of uncertainty.

Specific concerns from Panel members are as follows:

a. Group application of a UF_A (experimental animal to human extrapolation factor) is problematic if both human and animal data are used in cumulative risk assessment; there should be a method of adjustment which accounts for the possibility of important inter-species differences for the individual chemicals in the CAG.

b. One member strongly disagreed with use of a UF_D for a CAG as a whole as this may not be satisfactory when dose-response relationships are not well-established for all chemicals in a CAG; the contribution of each chemical to uncertainty associated with the CAG is a function of relative exposure; it would be most practical to apply UF_Ds to individual chemicals prior to cumulative hazard assessment as the adjustments that would be necessary, based on exposure factors, from one cumulative risk assessment to another would likely be unworkable.

c. Serious critical evaluation of the use of Group Uncertainty Factors as put forward in the proposed guidance will depend on its application to real world examples.

d. One member suggested the proposed guidance be more explicit in its discussion of how uncertainty propagates through cumulative assessment.

e. In response to the third part of question 3, one member failed to find clear presentation of the strengths and weaknesses of the group UF approach in the proposed guidance.

f. One member suggested that additional thought be given to the issue of inter-species uncertainty factors and whether they should be applied at the outset individually (i.e. PoDs for individual compounds) or applied as a group after the fact. Difficulties arise when the database includes studies comprised of both human and animal data for some compounds, but only animal data for others. In cases where correction needs are determined and must be made by extrapolation from animal to human, it is not enough to know that there is an identical mechanism of action in both species. Metabolism can differ greatly across species, and even from individual to individual, on a compound by compound basis which is not readily predictable from analysis of chemical structure. Therefore, even when human data are available for most compounds, and these data demonstrate equivalent sensitivity between humans and rats, it may still be necessary to acknowledge that remaining compounds in the CMG might differ markedly across species. Differences in sensitivity should be accommodated either by adjusting PoDs for those compounds, or by applying a weighted correction factor "after the fact" to the whole group, in a manner reflective of relative risk (perhaps as determined by MOE for the individual compound as a fraction of the combined MOE).

Issue 4. Methods for Estimating the Cumulative Toxicity

As discussed in *Chapter 5.6* of the background document, one of the steps in the cumulative risk assessment process will be to select a method to cumulate dose or exposures. This method will serve to normalize differences in the toxic potencies among the chemicals in the cumulative assessment. Precedence in the Agency's 1986 and revised 1999 *Guidance for Conducting Health Risk Assessment of Chemical Mixtures* describes several techniques for estimating risk to multiple chemicals. The cumulative guidance focuses on the component-based dose addition methods used in the Agency's chemical mixture assessment guidance document. Two methods, a margin-of-exposure approach and an approach using relative potency factors, are presented.

4a. Does the SAP agree that both methods are valid to consider for estimating cumulative risk associated with exposures to chemical that cause a common toxic effect by a common mechanism? Has the document clearly described these two approaches and their strengths and limitations? Are there other methods that OPP should consider?

4b. It is anticipated that most mechanisms of toxicity encountered currently will be nonlinear dose-response relationships. Nevertheless, for mechanisms of toxicity consistent with linear dose-response relationships, does the SAP agree that using the relative potency factor approach by summing the slopes of the dose-response curves is an appropriate method? If not, what methods would the SAP recommend for low-dose linear extrapolations of risk?

The members of the Panel concluded that the Margin of Exposure (MOE) method and the Relative Potency Factor (RPF) method are clearly useful methods to apply in initial investigations of cumulative risk assessment. The validity of these methods remains to be determined. Since both methods are based on dose-addition, they are conceptually valid only when the dose-

response relationships for the individual chemicals are parallel. If the relative potencies are based on the point of departure (POD), then the relative potency method and the MOE method are the same. As long as the point of comparison for relative toxicity is of the same risk level in each compound (something NOAELs poorly approximate, for instance) and the dose-response curves are parallel, even relative potencies based on other comparisons besides the POD will give the same result as the MOE approach. However, some members of the Panel suggested the following:

a. The assumptions underlying the use of these methods should be explicit in the proposed guidance and should at least clarify 1) whether or not the dose-response curves for members of a CMG are assumed to be parallel and 2) if dose-response curves are assumed to be parallel, what is the impact for assessment methodology if the curves are not parallel?

b. Some effort should be expended to assess the accuracy of these two methods and the assumptions that underlie each method and to rigorously compare the methods; analysis of appropriate hypothetical data sets may help identify factors that could introduce significant error.

c. The proposed guidance should clarify if both of these methods will be applied to each cumulative risk assessment. If not, what criteria will be used to select one method over the other?

The proposed guidance does describe strengths and limitations of the two methods. There are actually two separable issues to consider in combining toxic responses by these methods: (1) calculation of relative potency among the agents in question (i.e., the fraction of the total risk coming from each agent) and (2) the absolute degree of potency (i.e., the total risk). If one acknowledges that each agent's individual toxicity is determined with experimental error, and that the degree of such error may vary among agents, then the two methods differ in their sensitivity to such error. The relative potency method incorporates any error in the determination of the absolute potency of the index chemical into the calculation of all the relative potencies, and so it is a good idea to use the best characterized agent as the index. The MOE method spreads the effects of error among all the agents, since no single one need be chosen as the index. This makes the result less vulnerable to the error in one particular agent's dose-response determination, but it also means that the overall error is from a combination of the better- and worse-estimated curves (in a sense, averaging the error over them) and not just of the curve thought to have the least error. Further discussion of the sensitivity of each method to experimental errors in determination of points-of-departure is warranted. Proportionality of differences in MOEs should not be interpreted as proportionality in differences in risk. The temptation to act as though an MOE 10-fold larger implies a risk 10-fold lower should be avoided. If dose-response curves are not estimated (e.g., if only NOAELs are used of if PoDs are presented without presenting the fitted curves they are derived from), then relative risks at different levels of exposure cannot readily be compared. Some difficult examples of cumulative risk assessment that require application of uncertainty factors might be informative for comparing strengths and limitations of the methods of MOE and RPF.

The Panel did not specifically propose other methods for the Agency to consider.

However, the Panel did ask for more information on methods that were rejected for cumulative risk assessment. One Panel member recognizes that the Agency needs to avoid two unsatisfactory extremes: 1) insisting on "ideal" methodology which cannot be used with available data; 2) accepting a "purely practical" methodology with poor scientific rationale, which offers no way to improve risk analyses when better data become available. The best compromise is to define methodology with reference to the data that are desired and can ultimately be provided. Such a methodology allows for the frequent cases when ideal data are lacking and analyses must use practical surrogates. For example, rather than base risk assessments on NOAELs, one could base them on benchmark doses, while allowing use of NOAELs when benchmark doses are not available. This strategy encourages the generation of better data, it allows policy choices to evolve as the science improves, and it puts less-than-ideal analyses in their proper perspective.

The Panel did accept EPA's proposal to use the RPF approach by summing the slopes of the dose-response curves when linearity has been established in dose-response relationships.

Issue 5. Case Study

In *Appendix A* of the guidance document is a case study on organophosphorous pesticides.

5. Does this case study provide a clear example of the application of the hazard and dose-response elements of the draft guidance?

The Panel agreed that the use of case studies to illustrate application of the proposed guidance to cumulative risk assessment is a valuable addition to the proposed guidance document. It was agreed that the straightforward case study provided for review gave a general feel for the Agency's weight-of-evidence approach; however, all Panel Members considered this example too simplistic. A more complex, realistic case illustrating application of uncertainty factors and considerations of sensitive subpopulations would would provide much better guidance. Panel members suggested the following:

a. Providing more than one case study should be considered.

b. Correlating case study examples with technical guidance chapters would allow the proposed guidance to illustrate points as they are being developed in the text.

c. The case study examples should clearly identify the portions of the cumulative risk assessment approach that are based on science and those portions that are based on policy.

In addition, one Panel member noted the following points concerning the case study:

a. The selection of data for NOAEL derivation is confusing. The narrative refers to studies

that are not listed in the summary tables, and the tables show studies that aren't addressed in the narrative. In at least one case, a NOAEL is selected at a dose which is an effect level in one species. No rationale for this is provided.

b. Section 1.3.11 of the background document ostensibly provides several reasons for selecting red blood cell ChEI as the recommended endpoint for use in the cumulative risk assessment model; instead, it presents reasons for selecting the rat data as most appropriate over data from other species. These are not the same issue, and this section needs to be clarified.

c. On pg. 83 of the background document, data are presented and interpreted as supporting a dermal bioavailability of 42%. The approach presented cannot be used for this purpose. Two doses, each of different size administered by different routes and giving a different response level, cannot be used to estimate bioavailability unless blood concentration-effect relationships are extraordinarily well understood. There is no indication that this is the case in this example. This dermal bioavailability assessment is not essential to the case study, and should be deleted.

d. In quantifying cumulative risk, the example states (e.g., pg. 92) that, " In practice, the risk data for each pathway would be combined over time and presented as a distribution." There are significant technical issues that need to be addressed and resolved before this can be accomplished. Additional detail on this aspect needs to be provided.

e. A Uf_D was not applied in this example. The rationale for this decision needs to be included in the example.

SAP Report No. 99-05E, November 18, 1999

REPORT:

FIFRA Scientific Advisory Panel Meeting, September 23, 1999, held at the Days Inn Hotel, Arlington, Virginia

Session V - A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding:

Review of American Cyanamid Company's Probabilistic Assessment for Chlorfenapyr and Request for Guidance on Problem Formulation

Mr. Paul I. Lewis
Designated Federal Official
FIFRA/Scientific Advisory Panel
Date:_____

Herbert Needleman, M.D.
Chair
FIFRA/Scientific Advisory Panel
Date:_____

Federal Insecticide, Fungicide, and Rodenticide Act
Scientific Advisory Panel Meeting
September 23, 1999

Session V: Review of American Cyanamid Company's Probabilistic Assessment for Chlorfenapyr and Request for Guidance on Problem Formulation

PARTICIPANTS

Chair
Herbert Needleman, M.D., Professor of Psychiatry and Pediatrics, University of Pittsburgh Pittsburgh, PA

FQPA Science Review Board Members
William Adams, Ph.D., Kennecott Copper Company, Magna, UT
George Cobb, Ph.D., Texas Tech University, Lubbock, TX
Michael Daniels, Sc.D., Assistant Professor of Statistics, Department of Statistics, Iowa State University, Ames, Iowa
Paul W. Eslinger, Ph.D., Staff Scientist, Pacific Northwest National Laboratory, Richland, WA
Scott Ferson, Ph.D., Applied Biomathematics, Setauket, NY
Christian Grue, Ph.D., Washington Cooperative Fish & Wildlife Research Unit, Seattle, WA
Owen Hoffman, Ph.D., President and Director, SENES Oak Ridge, Inc., Oak Ridge, TN
Ms. Linda Lyon, United States Fish and Wildlife Service, Arlington, VA
Dwayne Moore, Ph.D., The Cadmus Group, Ottawa, Ontario, Canada
Ken Portier, Ph.D., Associate Professor, Department of Statistics, University of Florida, Gainesville, FL
Bill Williams, Ph.D., Kennedy/Jenks, Portland, OR

Designated Federal Official
Mr. Paul Lewis, FIFRA Scientific Advisory Panel, Office of Prevention, Pesticides and Toxic Substances, Environmental Protection Agency, Washington, DC

PUBLIC COMMENTS

Oral statements were received from:
American Cyanamid Company

Written statements were received from:
James Gagne, Ph.D., American Cyanamid Company
Mr. Damian Preziosi and Mr. Dan Woltering, The Wineberg Group
Geoffery Hill, Ph.D, Auburn University
Ms. Kelly Tucker, American Bird Conservancy
Mary Henry, Ph.D., US Fish and Wildlife Service

INTRODUCTION

The Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Scientific Advisory Panel (SAP) has completed its review of the set of scientific issues being considered by the Agency regarding review of American Cyanamid Company's probabilistic assessment for chlorfenapyr and request for guidance on problem formulation. Advance public notice of the meeting was published in the Federal Register on September 3, 1999. The review was conducted in an open Panel meeting held in Arlington, VA, on September 23, 1999. The meeting was chaired by Herbert Needleman, M.D. Mr. Paul Lewis served as the Designated Federal Official.

In December 1994, the Agency received a request for registration for the use of the pyrrole insecticide, chlorfenapyr on cotton. As a part of the registration package, American Cyanamid Company submitted a probabilistic assessment. The Agency sought FIFRA SAP input regarding: 1) American Cyanamid's probabilistic assessment and how it may be improved, 2) the Agency's review of the assessment, and 3) the utility of using this assessment to characterize the risk of chlorfenapyr use on cotton to birds in cotton agroenvironments. In addition, the Agency sought guidance regarding problem formulation. Ms. Denise Keehner (Office of Pesticide Programs, EPA) opened the session discussing the objectives of the Agency's presentation. Edward Odenkirchen, Ph.D. (Office of Pesticide Programs, EPA) provided an overview of chlorfenapyr chemical characteristics and properties. Mr. Jim Jones (Office of Pesticide Programs, EPA) identified major issues for inclusion in the ecological risk assessment for chlorfenapyr. Ms. Denise Keehner (Office of Pesticide Programs, EPA) completed the Agency's presentation reviewing questions posed to the Panel.

CHARGE

The specific issues to be addressed by the Panel are keyed to the background document and are presented as follows:

1. American Cyanamid Company's probabilistic risk assessment defines a bird population as " *a group of individuals belonging to the same species inhabiting the southern and western U.S. where cotton agroenvironments exist. Natural populations of birds are not geographically*

bound to arbitrary spatial areas, nor can they realistically be assigned temporal boundaries." (Avian Probabilistic Ecological Risk Analysis for Chlorfenapyr (AC 303630) in Cotton. Submitted by American Cyanamid Company to the Agency April 22, 1999. MRID No. 448098-01. Study completion date: December 9, 1998.)

a. Does the SAP agree that populations of concerned bird species associated with cotton agroenvironments are not geographically bound to spatial areas?

b. Since the Agency is concerned with how birds on cotton fields treated with chlorfenapyr are affected, would a scale by acre or field be more appropriate?

c. Can the SAP suggest further refinements in the scale of probabilistic assessments?

d. Are there important ecological impacts not accounted for in using the population definition and geographic scales assumed in the probabilistic risk assessment?

e. If so, are there sufficient data to address these effects at smaller geographical scales for more limited definitions of populations?

2. Given the scale established for American Cyanamid Company's probabilistic risk assessment, are the data used in the probabilistic risk assessment sufficient in geographic scope and technical rigor to establish the probability distributions for the wildlife food residue and dietary proportions (field and buffer) assumed in the probabilistic risk assessment with reasonable scientific certainty? If not, would a smaller geographical scale be more appropriate for use with the available data, and what scale would be recommended?

3. Based on the registrant's definition of population, American Cyanamid's probabilistic risk assessment concludes that the "*results of the present probabilistic risk analysis demonstrate that the risks to birds from recommended use of chlorfenapyr applied to cotton are negligible.*" Does the SAP believe that this conclusion can be reasonably extrapolated from the registrant's probabilistic risk assessment to populations and geographic scales intermediate to nationwide and individual treated fields? If not, can the SAP recommend appropriate modifications to the assessment approach and any additional data requirements to facilitate such an extrapolation?

4. Referring to EFED's evaluation of the registrant's probabilistic risk assessment ("Evaluation of 'Avian Probabilistic Ecological Risk Analysis for Chlorfenapyr (AC 303630) in Cotton. MRID No. 448098-01 ' by Ed Odenkirchen and Alex Clem, EFED) and the exposure model used by the

registrant:

a. Are EFED's concerns regarding the modification of exposure distributions through the use of the variables $P_{CAforage}$ and $P_{treatment}$ reasonable? Can the SAP suggest other quantitative methods of assessing population risks from reproductive effects in birds using treated cotton agroenvironments in the context of larger populations that would afford less uncertainty?

b. Does the SAP agree with the Agency's concerns regarding the variable P_{infest}?

c. American Cyanamid's probabilistic risk assessment relies on avian census data from a series of field studies described in MRID No. 444642-02. These data will be forwarded at a later date and are presented as being illustrative of avian use inside and outside of cotton agroenvironments (cotton field and buffer as defined in the probabilistic risk assessment). EFED's review has suggested that the nature of the study plots established in these field studies results in census data limited only to questions of avian use of in-field and buffer areas (i.e. the cotton agroenvironment only) and are not applicable to questions of avian use outside of cotton agroenvironments as defined by American Cyanamid's probabilistic risk assessment. Does the SAP agree with the EFED position?

d. Given that American Cyanamid's probabilistic risk assessment is concerned with risk quotients for time-zero (after application) residues, does the SAP agree with EFED that the approach for selecting peak values underestimates residues because of the total mass applied normalization of the third applications?

5. Referring to the discussions of probabilistic approaches from the July 22nd SAP:

a. Does American Cyanamid's probabilistic risk assessment attempt to quantify uncertainty associated with interspecies extrapolations of avian reproduction toxicity endpoints?. Given that the majority of birds observed in cotton agroenvironments are passerines, can the SAP propose an appropriate extrapolation uncertainty factor to account for differences in sensitivity between the tested mallards and quail to the untested passerines?

b. The SAP discussed approaches for investigation dependancy or covariance between variables. The registrant's response to questions regarding testing for such confounding factors is presented in the document included in this package entitled "CHLORFENAPYR: Avian Probabilistic Ecological Risk Analysis for Chlorfenapyr in Cotton: Response to Questions by EFED Reviewers". Given that a number of residue studies for different wildlife food items were conducted on different fields and that there appears not to have been plot-pairing of food item analyses for those studies with multiple food items investigated, does the SAP believe that investigations of covariance or dependancy between variables conducted for the registrant's probabilistic risk assessment are adequate?

6. Based on Mr. Jim Jones' presentation on the problem formulation for chlorfenapyr use on cotton, do you have any guidance and suggestions for improving the problem formulation?

PANEL RECOMMENDATION

The SAP does not believe that multiplying field level exposures by the ratio of fields sprayed with chlorfenapyr to total cotton fields as a means of estimating population level effects is valid. Clearly, risks at the level of fields treated with chlorfenapyr must be estimated as part of this risk assessment. A spatially explicit metapopulation model could then be used to determine the consequences of the field level risks for populations spread over large spatial areas.

The consensus of the Panel is that American Cyanamid's conclusion of negligible risk is not supported for any geographic scale. The scaling up of field exposures to the Cotton Belt scale was done incorrectly. The probabilistic exposure assessment had several other flaws that require remediation.

While the Agency and American Cyanamid Company should be commended for their efforts, the Panel urges increased cooperation and communication between the Agency and registrant for this and other risk assessments. Such effort is needed because both the registrant and the Agency are advancing the science while at the same time trying to meet regulatory requirements and/or make regulatory decisions.

DETAILED RESPONSE TO THE CHARGE

1. American Cyanamid Company's probabilistic risk assessment defines a bird population as *" a group of individuals belonging to the same species inhabiting the southern and western U.S. where cotton agroenvironments exist. Natural populations of birds are not geographically bound to arbitrary spatial areas, nor can they be realistically be assigned temporal boundaries."* (Avian Probabilistic Ecological Risk Analysis for Chlorfenapyr (AC 303630) in Cotton. Submitted by American Cyanamid Company to the Agency April 22, 1999. MRID No. 448098-01. Study completion date: December 9, 1998.)

a. Does the SAP agree that populations of concerned bird species associated with cotton agroenvironments are not geographically bound to spatial areas?

While the question is somewhat vague, the Panel agrees with the underlying premise that political or economic regions are inappropriate boundaries. Many species will be bound to habitats that only occur in subregions of the cotton belt (e.g., the Arizona desert).

For the purposes of the risk assessment, bird species potentially exposed to insecticides are spatially bounded by the areas that receive pesticide application, spray drift, or runoff. It is important to note that runoff scenarios and species potentially exposed through this route are not considered. The issue is not bounded to a particular geographic area, but specifically foraging and nesting behavior differences in different eco-types that would impact the probabilities of selecting treated fields as well as the consumption patterns when in the fields. For example, return rates range considerably: 33% of adult female blue grosbeak return to a field while 45% of males

return.

There are localized populations that should have adapted foraging strategies and other habits that in effect mean for exposure assessment purposes, that a bird is not really free to forage elsewhere. Incorporating mean habits will not identify the organisms at highest risk.

Determining which species may be exposed to the pesticide must be based on intensive survey work. A species that spends relatively little time in and around the cotton field could be undetected in a survey but may still be subject to significant pesticide exposure. This is illustrated by findings of Tacha et al. (1994).

b. Since the Agency is concerned with how birds on cotton fields treated with chlorfenapyr are affected, would a scale by acre or field be more appropriate?

The issue of appropriate scale is both a science and policy issue. For example, if legislation and policy dictated that protection of all bird individuals is the objective, then an acre or field scale may be appropriate. If protection of populations is the objective, then a larger spatial scale would be required. It is not clear whether the Agency is solely concerned with effects at the population and higher levels of organization. For example, if chlorfenapyr had caused large bird kills where it was applied, then it might be possible that the Agency would take action to prevent future kills, even if the initial kills had no ramifications at the population level.

As was noted in earlier discussions of the FIFRA SAP, this type of analysis should be conducted at multiple scales, including biological (individual, population), geographic (field, multiple local fields [e.g., crop reporting district], county, region, and nation), and temporal (seasonal, annual, multiple years). Direction for scaling issues/needs should come from and be provided in the "Problem Formulation" step. Reasons for multiple scales include: the sporadic and most often localized nature of the target pest infestations and outbreaks; potential changes in "market share" depending on efficacy and treatment alternatives; inherent variability in associated bird species and communities and the target and non-target habitats that support them; the fact that much of the necessary data are available; and that this type of analyses lends itself to differences in the exposure scenarios related to these scales and the selection of the most appropriate "focal" species for the various analyses. The Panel strongly argues that scale is not just a risk management issue, but is also an ecological one, i.e.: Are there labeling/mitigation options that can address local, regional and/or national concerns? Subsequent analyses by a Panel member in which some of the input variables were modified indicates that local and regional impacts of the chemical on individual birds and local populations may be severe.

Fields are more appropriate that include edge habitat. Risks by acre are not true representations of exposure potential since most species nest in edge habitat and may forage in the planted fields. This dynamic cannot be accounted for if assessing by the acre or if evaluating based on average concentrations across entire fields and edge. The probability of chlorfenapyr (or any other pesticide) occurrence across treated fields and border areas must be considered through probability density functions.

Even if American Cyanamid and EPA were to agree that the *de minimus* level of concern was for populations over a large spatial scale, the approach used by American Cyanamid to assess risks to bird populations was inappropriate. In essence, their approach was to estimate field level exposures and then multiply these exposures by the probability that a field will become infested and the probability that an infested field will be treated with chlorfenapyr. Not surprisingly, the resulting exposure distributions for different guild representatives turn out to be well below the effects threshold.

Imagine that each Latin Hypercube sample is one individual out of a population of 10,000 red-winged blackbirds. Given the American Cyanamid results, one would conclude that only a few of the 10,000 blackbirds randomly chosen from cotton fields in the Cotton Belt would be exposed to levels high enough to cause an adverse effect. Is that really the case? One Panel member re-ran the Cyanamid analysis, except that the probability of infestation and probability of being treated with chlorfenapyr variables was removed. Essentially, this estimated the probability of exceeding the effects threshold for birds foraging in and near fields treated with chlorfenapyr. The result was that there was a 57% probability of red-winged blackbirds being exposed to chlorfenapyr levels above the effects threshold in fields treated with chlorfenapyr. If all cotton fields were treated with chlorfenapyr, 5716 individuals would be exposed to chlorfenapyr levels above the effects threshold. This is not the case, however, because not all fields are infested and not all infested fields are treated with chlorfenapyr. Multiplying 5716 by the means developed by American Cyanamid for proportion of infested fields (0.098) and the proportion of infested fields treated with chlorfenapyr (0.2) would indicate that 112 individuals out of the population of 10,000 blackbirds would be exposed at levels above the effects threshold, an answer which is far above the American Cyanamid estimate. Whether effects to 112 individuals out of a population of 10,000 constitutes a population level effect is a legitimate question, the answer to which would depend on spatial distribution of affected individuals, their demographic characteristics, other stressors impinging on the population, etc.

The above analysis clearly indicates the fallacy of multiplying field level exposures by the ratio of fields sprayed with chlorfenapyr to total cotton fields as a means of estimating population level effects. Clearly, risks at the level of fields treated with chlorfenapyr must be estimated as part of this risk assessment. A spatially explicit metapopulation model could then be used to determine the consequences of the field level risks for populations spread over large spatial areas.

Most members of the Panel believe that asking about population-level effects makes sense only when effects have been observed at the individual level. They suggest that not observing significant individual-level effects should be considered evidence that no effects could be manifested at the population level. In effect, this uses an individual-level assessment as a screening assessment. This may not always be prudent however. Effects at the lower level of aggregation may be harder to demonstrate statistically, for reasons involving the details of sampling, than effects at the population level. Even statistically negligible differences should be propagated through a population model to discern whether they might be important at the higher level. Furthermore, apparent modest impacts on mortality or reproduction may develop over time into substantial consequences for the population as a whole. Depending on the nature of density

dependence in the species in question, seemingly small consequences for individual organisms can sometimes be magnified to major consequences for populations. Such magnification of impacts at the population level is a phenomenon well known in population ecology (especially among threatened and endangered species) and should not be summarily discounted.

Distributions for input should be consistent with the scale of the assessment. Therefore, the frequency distribution of raw data may not be appropriate for risks that are summarized over large scales of time and space.

One Panel member commented that temporal scales become more important as geographic scales shrink. Thus, reducing the geographic scale may actually increase data needs rather than decrease them as one might expect.

c. Can the SAP suggest further refinements in the scale of probabilistic assessments?

As in any ecological risk assessment, the appropriate spatial and temporal scales for the chlorfenapyr assessment depend on the risk management goals (derived in part from legislation and policy), question being asked, stressor properties, use patterns, fate, and receptor characteristics. A well thought out problem formulation is required before any probabilistic risk assessment is carried out.

Assessing population level risks posed by chlorfenapyr for multiple species over an area as large as the Cotton Belt would be a very difficult task. As an alternative, the Panel suggested the use of a tiered probabilistic approach. A more manageable first step would be to use the problem formulation exercise to identify the areas that are likely to have the highest levels of chlorfenapyr (e.g., because environmental conditions are not conducive to chlorfenapyr degradation) and the receptor species that will have the highest exposures because of their foraging behavior in cotton agroenvironments. Alternatively, several areas with low, medium, and high chlorfenapyr exposures could be selected so that the range of risks at the field level can be investigated. The next step would be to select one to several fields for detailed probabilistic assessment. If the risks are negligible for highly exposed bird individuals in and near fields with the highest levels of chlorfenapyr, then there may be no need to consider larger spatial scales or higher levels of organization. Hierarchy theory (O'Neill et al. 1986; Moore 1998) suggests there can be no effects at the population level unless there are effects at the individual level. Therefore, an assessment may only need to demonstrate there are no adverse consequences at the lower level (although care must be taken to ensure there is enough statistical power to detect effects among individuals if they are present, as presented in the previous question). The advantages of starting at the field level are that the risk model is much simpler and data gaps are more easily filled. If risks are serious at the field level, or if data are inadequate to show they are not serious, then the assessment needs to move to a higher tier that perhaps involves the use of a stochastic metapopulation model.

The probabilistic assessment is flawed and its conclusion is unjustified, but the problem is not really the seemingly strange definition of ‘population.’ It may be reasonable to define a

population of birds extending throughout the Cotton Belt, although most ecologists would probably used the term 'metapopulation' for this aggregation. The real problem is that the assessment has attempted to scale the impact to the population level in a simplistic way that does not make biological sense. The effect of interest may be at the population level, but the impacts are on individual birds. It is individuals that experience mortality or fail to reproduce. The calculations have inappropriately diluted potential local impacts by averaging over the entire region, producing obviously underestimated risk quotients. Contrary to the conclusions of American Cyanamid, this argument is insufficient to pass even a screening assessment.

Given the interest by risk managers in the potential effects on local populations and larger-scale metapopulations, the assessment should be expressed at these higher levels of biological organization. But one cannot simply scale the results from a traditional risk calculation, which essentially focuses on individuals. There must be an explicit extrapolation to the population level, such as by means of Leslie-like matrix models or individual-based population models. Stochastic versions of such models that would be appropriate for this extrapolation in a probabilistic risk assessment have been widely used in ecological risk assessments over the last decade or two. Several convenient software tools are available for such assessments. These tools take demographic information about the species of interest such as survival, fecundity and dispersal rates, temporal variability in those rates, and density dependence phenomena to produce population-level summaries of ecological risk. A geographical, spatially explicit demographic model would be needed to propagate the potential impacts to the level of an entire metapopulation extending throughout the Cotton Belt. Such metapopulation models have been used in several ecological risk assessments over the last ten years.

It is possible, within the context of a risk assessment, to project potential ecological effects to the level of a local population, a metapopulation, or even a regional population. Although the data needs for these projections are of course greater than for an assessment of individual-level effects, useful and appropriate information is probably available for many species of concern. Moreover, so long as an appropriate uncertainty analysis accompanies the demographic projection, whatever data are available are *sufficient* for a risk analysis because the assessment thereby includes a statement of its own reliability.

d. Are there important ecological impacts not accounted for in using the population definition and geographic scales assumed in the probabilistic risk assessment?

Population definition and geographic scale at which the assessment is to be performed are critical components of a risk assessment that are best defined at the beginning of the assessment. It is important that key stakeholders involved with the assessment reach consensus on these important points. There are no absolutes in the definition of population or the scale to be assessed. EPA Ecological Risk Assessment Guidelines recommend that risk assessments be performed at the population and not the individual level. However, there is also recognition that individuals can be important if the species is threatened or endangered. What is really critical in this process is that risk is determined and explained at a level (species/population) and scale (local/regional) that is relevant to the chemical of interest and relevant to the interest of the

stakeholders. One approach is to evaluate risk at more than one scale so the magnitude of risk both to the population (broad scale) and to individuals (local scale) is known and properly communicated in terms that are easily understood by all parties. Reduction in the scale often means that the risk to individuals goes up, but the number of individuals likely to be impacted goes down, hence there is a balance in the assessment. A regional reproductive risk may be 2% whereas the local risk may be 20%. Using either value as the single most important expression of avian reproductive risk would be an inappropriate conclusion. This balance needs to be presented and evaluated by the risk manager in the overall decision concerning the use of a given pesticide.

Specific to the question, if the level of interest is the population and the scale of interest is regional, and all the appropriate data are collected on exposure across the scale to the species of interest and the effects are well characterized on all the species of interest, there should be no concern that important ecological impacts are not accounted for. Risk is the summary of all probabilities of effects (for a given species) across the scale of interest. What might be at question with the current risk assessment is whether or not sufficient detail is known about exposure to multiple species both in treated and non-treated cotton fields and buffer areas to properly aggregate risk across scale. Additionally, the mallard duck reproductive study is used as a surrogate for all passerine birds. This is likely over-protective for some birds and under-protective for others. Concern for the existing assessment might arise from general assumptions about the feeding behavior and dietary exposure used for bird groupings. The risk assessment could be improved with greater detail in the exposure estimates.

In summary, conducting an assessment at a large scale doesn't by definition mean that important ecological impacts are not accounted for, but the potential does exist because regional scale assessment may lack local detail. For example, sub-populations within a species that have optimized foraging habits for a given geographic region and crop type likely have different exposure potential due to differential behaviors, including foraging strategies. The question then becomes one of risk management as to whether or not local scale effects, if observed, are important ecological impacts.

e. If so, are there sufficient data to address these effects at smaller geographical scales for more limited definitions of populations?

Risk analysts, like scientists in general, are acutely aware of their data limitations. Risk analysts do not, however, have the luxury of declining to answer until enough information becomes available. Indeed, if there were enough data to satisfy the analyst, there would scarcely be any need for an analysis in the first place: the answer would surely already be clear.

The critical data needs for an assessment of chlorfenapyr risks to birds are information on residues in prey items and knowledge of foraging behavior in cotton agroenvironments. With regard to residues, American Cyanamid has generated some data for weed seeds and insects in treated fields. The data, however, are spotty for buffer areas and for fruits. The only high quality study on foraging behavior was the one done as presented by Dr. Geoffrey Hill (Auburn University) during the public comment section of this FIFRA SAP meeting. The relevance of the

existing data to areas without measured data is unknown. If the assessment were to be focused to a small number of fields, it would be feasible to collect the required data. Alternatively, stochastic models could be used to estimate residue levels in treated fields and in surrounding buffer areas. The work of the Spray Drift Task Force is especially relevant for estimating levels in buffer areas. These models could be calibrated using the existing residue data.

As stated by the Agency, insect residue data were collected from only one field study and then were extrapolated to all agroenvironments. This may be legitimate, but more locations should be studied and statistical tests for homogeneity performed. Otherwise, the analysis should account for the considerable uncertainty in extrapolating from just one site. If residues do differ by location, collecting information about locations might support development of a model for how residues depend on various characteristics of a location. The fruit residue data were collected on fruits that are not good surrogates for fruit consumed by birds. Data should be collected on relevant fruits and appropriate correlation studies performed. The data on weed seed residues have the same problem as the insect residue; data collected at very few sites (in this case, two).

Detailed data on species diversity and habits in vast agroecosystems exist within the data bases of several agrochemical corporations. But these data are not often shared. More data or data sharing would be necessary to perform probabilistic assessments that incorporate differential behaviors of sub-populations.

Uncertainty analysis is the practical means by which analysts determine whether data are sufficient for decision purposes. Uncertainty analysis assesses the reliability of the values and conclusions reached in a risk assessment. In the interplay between the location and uncertainty of a value, there are three possible outcomes, sometimes described as the ‘good’, the ‘bad’ and the ‘ugly’. Good outcomes are those in which the value is clearly within a tolerable range. Bad outcomes are those in which the value clearly exceeds tolerable levels. Ugly outcomes are those in which the uncertainty around the value precludes the clear classification as either tolerable or intolerable. The most important ability of such analyses is to distinguish ugly results from the other two cases. Of course, more data would always be *desirable*, but they may not be *essential* for making a decision. In some cases, when the uncertainty about an answer is small relative to its magnitude, the appropriate decision is clear. A comprehensive uncertainty analysis will make such cases evident, and it will also indicate when more data are needed to make a reliable decision (and will typically be very helpful in directing the requisite empirical effort). Using uncertainty analysis means that there are always enough data for an assessment to be useful.

Thus, an uncertainty analysis performed for a probabilistic assessment provides an indication of the quality of the results. There are enough data to perform the initial assessment. However, since the initial assessment apparently indicates the possibility for significant avian impacts from the use of chlorfenapyr, more data are needed to resolve more precisely the level of potential impacts.

2. Given the scale established for American Cyanamid Company's probabilistic risk

assessment, are the data used in the probabilistic risk assessment sufficient in geographic scope and technical rigor to establish the probability distributions for the wildlife food residue and dietary proportions (field and buffer) assumed in the probabilistic risk assessment with reasonable scientific certainty? If not, would a smaller geographical scale be more appropriate for use with the available data, and what scale would be recommended?

As stated by the Agency, insect residue data were collected from only one field study and then were extrapolated to all agroenvironments. This may be legitimate, but either more locations should be examined and homogeneity explicitly tested (statistically) or the analysis should account for this considerable extrapolation uncertainty. If residues differ by locations, collecting information about locations might support development of a model for how residues depend on various characteristics of the location, i.e., explain heterogeneity. The fruit residue data are measurements on fruit that may not be good surrogates for fruit consumed by birds. Data should be collected on the relevant fruits and appropriate correlation studies performed. A similar problem exists for weed seed residues - data are generalized from two sites to all cotton agroenvironments across the U.S. Statistical tests for homogeneity should be performed. If the data are not homogeneous, but heterogeneity can be explained by covariates, one could then estimate data for whole U.S.

In summary, the existing data are not sufficient in geographic scope and technical rigor for the large-scale analysis performed by American Cyanamid. Further, the data may be too limited for small analysis (broader geographic), which goes beyond the actual site investigated. Avian diet and residue studies need to be replicated in other cotton regions if a smaller scale analysis is to be performed.

3. Based on the registrant's definition of population, American Cyanamid's probabilistic risk assessment concludes that the *"results of the present probabilistic risk analysis demonstrate that the risks to birds from recommended use of chlorfenapyr applied to cotton are negligible."* Does the SAP believe that this conclusion can be reasonably extrapolated from the registrant's probabilistic risk assessment to populations and geographic scales intermediate to nationwide and individual treated fields? If not, can the SAP recommend appropriate modifications to the assessment approach and any additional data requirements to facilitate such an extrapolation?

The interpretation of the output distribution of the probabilistic risk analysis is not well defined. Is the distribution the variability of exposure in individual birds or is it uncertainty in exposure for the average bird in a defined sub-population (guild)? The data seemed aimed at average birds (which do not actually exist), and the risks to individual birds does not seem to be addressed. The probability of a bird selecting a cotton field does not include the value of 1 for a bird selecting a foraging area; the cotton field and surrounding habitat may be the only available area. There are foraging areas that are defended by dominant individuals and this may alter exposures to other individuals. By incorporating these biased-low distributions for selecting a cotton agroecosystem to forage and of selecting actual treated field versus edge habitat also

lowers the likelihood that the model will show the highest exposures actually observed in the field.

American Cyanamid's conclusion of negligible risk is not supported for any geographic scale. The scaling up of field exposures to the Cotton Belt scale was done incorrectly as noted in the Panel's response to question 1(b) above. Even ignoring the dilution that arises by multiplying exposure by the proportion of fields treated with chlorfenapyr, the probabilistic exposure assessment had many flaws. These include:

Using concentrations instead of doses. Clearly, Carolina Wrens have substantially higher food intakes per kg biomass than do mallards. By using concentrations in the quotient numerator (wren exposure) and denominator (mallard NOEC), this important contributing factor to species sensitivity differences is overlooked.

Using quotients instead of risks. Quotients have very little ecological meaning to risk managers. Risks should have been expressed as probability and magnitude of effect curves. For chlorfenapyr, a dose-response curve may need to be generated for reproductive effects of species of interest in order to overcome this flaw.

Dependencies between input variables were ignored. The correlation matrix for the inputs used in the exposure equation indicates quite strong correlations ($r>0.6$) between many of the residue variables. These dependencies could have a significant effect on the shape and breadth of the resulting exposure curve.

There were statistical errors in the analysis. The American Cyanamid probabilistic risk assessment had two variables for which the geometric standard deviation was defined as being less than one. This is not possible. The correlation matrix also was not positive, semi-definite. For example, seed and fruit residues in the field were positively correlated, seed residues in field and buffer were positively correlated, but fruit residues in the field and seed residues in the buffer were negatively correlated. Again, this does not make sense.

The choice of some distributions is not well supported. The variables for proportion of fields infested, proportion of birds foraging in the cotton agroenvironment, etc. were assigned lognormal distributions. For proportion variables that scale from 0 to 1, a distribution such as the beta distribution would be more appropriate.

The parameterization of input distributions is not well thought out. In the American Cyanamid analysis, lognormal distributions were fit to measured residues data. However, any bird feeding in a field over time will, to some extent, average their intakes over time (sometimes consuming more contaminated food items, other times less). Therefore, the residue distributions should represent our uncertainties about what this averaged exposure would be, not simply the scatter of residue concentrations in a series of fields.

Not all cotton agroenvironments are the same. American Cyanamid's analysis ignored the

likelihood that risks vary between habitats and ecoregions. Some cotton fields are in forested areas, others in deserts, etc — why would risks be the same for such disparate habitats?

Important bird behaviors and chlorfenapyr characteristics were not considered. Birds are not likely to select a cotton agroenvironment and then stay there for long periods of time. More likely individuals will fly in and out of cotton field areas. This feature could be explicitly modeled. Similarly, chlorfenapyr degradation over time could be explicitly incorporated in the exposure model. Other pathways such as dermal contact with feet that could be considered since other pesticides display dermal doses equivalent to dose via ingestion. These flaws make the American Cyanamid conclusion of negligible risks very tenuous.

If the model were appropriately formulated, alternative intermediate geographic scales could be considered without much additional work. An alternative question which could be posed and answered is, What is the risk to birds only at cotton fields (including their appropriate buffer zones) where chlorfenapyr is used? This is a much narrower question and focuses the question only on the issue of chlorfenapyr at specific sites. This approach leaves open the questions about what the boundaries are for a population and focuses on risk across treated sites. The data in the present risk assessment should contain much of the data needed for this recalculation.

4. Referring to EFED's evaluation of the registrant's probabilistic risk assessment ("Evaluation of 'Avian Probabilistic Ecological Risk Analysis for Chlorfenapyr (AC 303630) in Cotton. MRID No. 448098-01 ' by Ed Odenkirchen and Alex Clem, EFED) and the exposure model used by the registrant:

a. Are EFED's concerns regarding the modification of exposure distributions through the use of the variables $P_{CAforage}$ and $P_{treatment}$ reasonable? Can the SAP suggest other quantitative methods of assessing population risks from reproductive effects in birds using treated cotton agroenvironments in the context of larger populations that would afford less uncertainty?

Use of these dilution factors is completely inappropriate in a situation where individual birds are exposed to a pesticide. These factors average the dietary exposure across the entire foraging population. If this scenario were true there would be no reason for probabilistic risk assessment. A good example of the fallacy of this approach can be found in the current arsenic poisoning situation in India. If risk assessors were to evaluate the arsenic concentration in all ingested water in the Indian subcontinent, the risks would appear non-existent. However, the exposed population is currently being decimated in the region of high arsenic exposure. The registrant's own risk assessment rationale argues against the inclusion of these parameters in the risk assessment. If there is a finite probability of a bird choosing a chlorfenapyr treated field, then the bird that does not select a chlorfenapyr treated field should be removed from the exposure assessment. The exposure of other birds to chlorfenapyr should not be ascribed to the unexposed birds.

One Panel member commented that it is a short-sighted ecological risk policy to evaluate ecological effects of individual pesticides in individual crop usages. For example, if there are 10 insecticides or 10 different insecticide usages that each have a 2% population impact, the aggregate effect will be approximately 19% population reduction. However, in ten separate evaluation processes, none of the registrations for insecticides or insecticide usages would be questioned. The median chlorfenapyr exposure exceeds 2% probability of effect in two of four modeled avian feeding guilds. Having presented this argument against evaluation for individual uses or for individual pesticides, it would seem prudent to perform a relative risk analysis of chlorfenapyr and other insecticides that are commonly used to control army worms.

One problem with the method used to estimate $P_{CAforage}$ is that the data comes from the bird activity analysis of the small point-center field edge counts. These same data were used to estimate $P_{field\text{-}forage}$ and $P_{buffer\text{-}forage}$. Because the coverage of non-cotton ecosystem area was by definition quite small for each point-center sampling, the estimate of $P_{CAforage}$ is poor. There are also temporal and spatial aspects of $P_{CAforage}$ which are not addressed. It is not clear from the data provided whether or not the sightings data under estimates or overestimates overall usage of agro-environments by birds. In summary, more site specific data at the local scale would be of interest.

Other Panel members supported the inclusion of $P_{treatment}$ into the overall risk assessment, but then commented that it is appropriate to consider risk both with and without the inclusion of non-treated fields in the assessment. In addition, it appears more appropriate to sum the acres infested with each pest for each year in deriving the extremes of the pest infestation estimates.

Concerns were expressed about the use of the mean values of the proportional census data because it doesn't capture the true temporal, geographic and species variability of field use versus agro-environmental use by birds from each cotton region where census data were collected. A better approach than just using the mean is to use the entire probability distribution. The probabilistic risk analysis should account for the variability in the diet per individual bird.

b. Does the SAP agree with the Agency's concerns regarding the variable Pinfest?

The Panel agrees with the Agency's concerns regarding the variable Pinfest. One Panel member provided a table showing that the approach taken by the registrant is in error (Table 1). The total fraction of infested acreage for 1995 is greater than 1. This also shows the shortcoming of the approach of isolating treatments targeted at individual pest species. Averaging data produces deterministic results and should not be included in assessments that claim to be probabilistic. The fact that calculations indicate an infestation fraction higher than 1 demonstrates the problem. Furthermore, the fraction of cotton treated appears to be estimated on the basis of individual pest species. The total infested acres are significantly more than indicated in the log-normal distribution on pages 19 and 25 of MRID 448098-01.

Table 1. Fraction of Cotton Infested

Year	Budworm	Armyworm	Fall Armyworm	Sum	Fraction of cotton
1988	6715292	NA	NA	6715292	0.5405
1989	4683150	529996	170336	5383482	0.4333
1990	6063315	803975	319195	7186485	0.5784
1991	8218849	472000	383300	9074149	0.7303
1992	5817902	307950	484300	6610152	0.5320
1993	6560980	1722250	554900	8838130	0.7113
1994	7316227	391700	187800	7895727	0.6355
1995	9259951	2505272	689325	12454548	1.0024
1996	4683150	317252	235195	5235597	0.4214

Several Panel members indicated that their concern was that this variable should not have been used in the exposure equation.

c. American Cyanamid's probabilistic risk assessment relies on avian census data from a series of field studies described in MRID No. 444642-02. These data will be forwarded at a later date and are presented as being illustrative of avian use of inside and outside of cotton agroenvironments (cotton field and buffer as defined in the probabilistic risk assessment). EFED's review has suggested that the nature of the study plots established in these field studies results in census data limited only to questions of avian use of in-field and buffer areas (i.e. the cotton agroenvironment only) and are not applicable to questions of avian use outside of cotton agroenvironments as defined by American Cyanamid's probabilistic risk assessment. Does the SAP agree with the EFED position?

A problem with estimation of $P_{CAforage}$ is that the data come from bird activity analysis of small point-center field-edge counts. The same data were used to estimate $P_{field\text{-}forage}$ and $P_{buffer\text{-}forage}$. Because the coverage of non-cotton ecosystem area was by definition quite small for each point-center sampling, the estimate of $P_{CAforage}$ is poor. There are also temporal and spatial aspects of $P_{CAforage}$ which are not addressed. It is not clear from the data provided whether or not the sightings data underestimates or overestimates overall usage of agro-environments by birds. In summary, more site specific data at the local scale would be of interest.

Another Panel member commented that it is difficult to determine if the estimate of $P_{CAforage}$ is an underestimate or an overestimate. Without defining the relevant population, this parameter is unidentifiable.

d. Given that American Cyanamid's probabilistic risk assessment is concerned with risk quotients for time-zero (after application) residues, does the SAP agree with EFED that the approach for selecting peak values underestimates residues because of the total

mass applied normalization of the third applications?

To be conservative, the residual concentrations of pesticide on vegetation should be compared to each application rate. A clear relationship between the pesticide application rate and the resulting vegetation residue at a study site should be generated whenever possible. Although the relationship between application rate and vegetation residue may vary, a conservative approach to estimating the likely tissue residues for the assessment might include the maximum, minimum, and mean residues expected for each application rate. It is clear that the physical conditions at each site will greatly impact the vegetation pesticide residue and there may be no single relationship that accurately reflects expected vegetation residue under every application scenario. Use of the range of vegetation residue values provides a reasonable method of estimating the range of potential exposure scenarios and can be incorporated into the probabilistic method.

In addition, the residue rate from the first application should be used. However, there is a general concern about using 'adjusted rates'. The approach recommended has an implicit assumption that the relation between the application rate and the residue is linear: i.e., if you double the application rate, you double the amount of residue and if you halve the application rate, you halve the amount of residue. This assumption should be tested by measuring residues for several different application rates and fitting a curve to the data. If the curve is not a straight line (linear), we may be under- or over-estimating the actual residues. We should also carefully model the degradation rate to be able to determine the residues after a series of applications.

The non-linearity of the three applications (i.e., different residue estimates per lb of chlorfenapyr, which is related to the degradation rate) may support the above and a curve might be fit such that the residue does not increase linearly with application rate. In addition, the Agency should carefully model the degradation rate.

5. Referring to the discussions of probabilistic approaches from the July 22nd SAP:

a. Does American Cyanamid's probabilistic risk assessment attempt to quantify uncertainty associated with interspecies extrapolations of avian reproduction toxicity endpoints? Given that the majority of birds observed in cotton agroenvironments are passerines, can the SAP propose an appropriate extrapolation uncertainty factor to account for differences in sensitivity between the tested mallards and quail to the untested passerines?

The probabilistic risk assessment does **not** attempt to incorporate the uncertainty associated with interspecies extrapolations. But, insofar as regulators accept the contention by the registrant that mallard values are likely to be conservative values for the passerines, it seems unnecessary to require an additional uncertainty factor for this interspecific extrapolation.

In the July SAP report, two methods are discussed to extrapolate from the tested species to the untested passerines. The references for Baril et al. were not readily available. There are

problems with using the second approach in this context (Luttik and Aldenberg). First, the methods assume that we have a random sample of LD50 data from several species. In our case, there is information on the LD50 for non-passerines and we are interested in making an inference about passerines. This calls into question the above assumption and could result in inaccurate estimates if passerines and non-passerines react differently to chlorfenapyr. Second, to obtain a 'better' estimate of the standard deviation, data are pooled across toxicants. This may yield a more precise estimate, but the method relies on the assumption that the standard deviation of log(LD50) is the same across all toxicants. This method relies on very strong assumptions that should be questioned.

However, the SAP discourages the approach adopted in the assessment of divorcing toxicity from the probabilistic treatment used for the exposure. The registrant has, in this regard, followed precedents and advice from EPA, but a majority opinion of the SAP considers this strategy inappropriate and unnecessary. The SAP majority opinion is that NOEC and NOAEL values are extremely poor summarizations of the available toxicological data, the richness of which should be more fully incorporated into the risk assessment. Indeed, the focus on the distribution of risk quotients as the final endpoint is itself problematic and confusing. Risk quotients lack any easily discernible meaning. It would be far more useful to simply estimate the probability of impacts of various magnitudes (such as, for instance, percentage bird population declines). If risk quotients are used to assess potential risk to populations, then at the least, the range of values used for the comparison of exposure to toxicity (Toxicity Reference Values [TRV], including the LOAEL and NOAEL) which appear in the denominator of the ratio should be scrutinized and ranked for credibility and relation to the endpoints used in the assessment. Many of the TRVs used in risk assessment have no physiological or ecological relevance to the species being evaluated. These estimates would be much more interpretable and, consequently, vastly more practical, both for the initial regulatory decision making as well as for any subsequent risk monitoring and risk management.

b. The SAP discussed approaches for investigation dependancy or covariance between variables. The registrant's response to questions regarding testing for such confounding factors is presented in the document included in this package entitled "CHLORFENAPYR: Avian Probabilistic Ecological Risk Analysis for Chlorfenapyr in Cotton: Response to Questions by EFED Reviewers". Given that a number of residue studies for different wildlife food items were conducted on different fields and that there appears not to have been plot-pairing of food item analyses for those studies with multiple food items investigated, does the SAP believe that investigations of covariance or dependancy between variables conducted for the registrant's probabilistic risk assessment are adequate?

This type of analysis should be conducted at multiple scales, including biological (individual, population), geographic (field, multiple local fields [e.g., crop reporting district], county, region, and nation), and temporal (seasonal, annual, multiple years). Direction for scaling issues/needs should come from and be provided in the "Problem Formulation" step. Reasons for multiple scales include: the sporadic and most often localized nature of the target pest infestations

and outbreaks; potential changes in "market share" depending on efficacy and treatment alternatives; inherent variability in associated bird species and communities and the target and non-target habitats that support them; the fact that much of the necessary data are available; and that this type of analyses lends itself to differences in the exposure scenarios related to these scales and the selection of the most appropriate "focal" species for the various analyses. The Panel strongly argues that scale is not just a risk management issue, but is also an ecological one, i.e.: Are there labeling/mitigation options that can address local, regional and/or national concerns? Subsequent analyses by a Panel member in which some of the input variables were modified indicates that local and regional impacts of the chemical on individual birds and local populations may be severe.

The correlation matrix given in Table 3 of the responses to questions by Agency reviewers (Addendum to MRID No. 44809801, transmitted by American Cyanamid on 16 July 1999) should be checked for positive semi-definiteness. It contains mutually contradictory values that represent an impossible state of affairs for the true correlations among the variables. The matrix must have been incorrectly compiled from discordant data sets, because it could not have arisen from a real data set. No adequate explanation of how data were paired was provided in the American Cyanamid probabilistic risk assessment.

As noted in the answer to question 3, some of the correlation coefficients in the correlation matrix were high ($r>0.6$), yet were not statistically significant. This means that sample sizes must have been very small. With these small sample sizes, there is virtually no statistical power for the test of “no correlation”. Statistical power in this instance is the probability of determining there is a correlation from the data-based test when a correlation actually exists.

Unfortunately, with a small sample size, it is impossible to say anything definitive about correlations among input variables based on the sampled data alone. That these correlations are not observed to be statistically significant does not mean that actual correlations will not have a crucial impact of the risk assessment results but specific analysis is required to discount this possibility. Assuming independence when dependencies exist can grossly over- or underestimate the dispersion of the final result. The analysis should incorporate the observed dependencies into the assessment. At a minimum, the analysis should explore the effects of the observed pattern of correlations on the risk results. With a Latin Hypercube sampling strategy, this may require translating the reported Pearson product-moment correlation coefficients into Spearman rank correlation values (or perhaps simply assuming they can be equated).

It would be useful for the assessment to include a sensitivity analysis in which the correlations are varied within plausible ranges. Of course, sensitivity analyses are cumbersome and are not generally sufficient to make firm conclusions about the lack of important consequences of dependencies on the risk results. One reason is that non-linear dependencies that are poorly characterized by correlation coefficients are difficult or impossible to discern in a sensitivity analysis. Other techniques such as dependency bounds analysis could be used to make definitive conclusions about the potential impact of correlations and dependencies among the input variables on the risk results.

6. Based on Mr. Jim Jones' presentation on the problem formulation for chlorfenapyr use on cotton, do you have any guidance and suggestions for improving the problem formulation?

American Cyanamid and the Agency are to be commended for their efforts in developing clear and thorough briefing documents. It was very apparent that both organizations devoted a significant amount of time (and money) toward the analysis. Such effort is needed because both the registrant and the Agency are advancing the science while at the same time trying to meet regulatory requirements and/or make regulatory decisions. The Panel agrees that the chlorfenapyr risk assessment should move from a traditional deterministic approach to evaluating risk to a probabilistic one, and the associated issues of scale (biological, spatial and temporal). This is why increased cooperation and communication between the parties is critical to the success of these efforts, irrespective of the final regulatory decision with respect to chlorfenapyr. If we want the "best available science" while at the same time "advance the science", then effective cooperation and communication is essential. With respect to chlorfenapyr, it appears both cooperation and communication could have been better, and had this been the case, some of the difficulties associated with the registrant's probabilistic assessment could have been avoided, or at least addressed much earlier in the process. The Panel wishes to express support of the overall concept of using probabilistic risk assessment for regulatory decisions. The development of an appropriate (mutually agreed upon) "Problem Formulation" would have greatly facilitated the effort by both parties. Sensitivity analyses determined that only a few variables were driving the results of the registrant's analysis, yet there was little or no communication/discussion about how these variables were to be addressed in the subsequent analysis.

Clearly, the spatial scale needs to be defined, preferably through joint consultations among assessors, managers, and stakeholders. Because Mr. Jim Jones noted that risks at the field scale are an important consideration in decision making, an assessment could start there because an assessment at that scale is the most manageable. Further, if there are no effects to highly exposed bird individuals in a range of treated fields, then there is a very low likelihood that chlorfenapyr is causing adverse population level effects at large spatial scales.

It is not clear what the Agency considers a relevant population. It seems that they have provided little guidance on relevant scales between field and national. The deterministic risk assessment identified ecotypes. Would the Agency be satisfied if the analysis were performed on populations with characteristics tailored to ecotype?

The probabilistic risk analysis purports to simulate exposures for individual birds as a frequency distribution. However, some of the data seem to be developed for the average of all birds. The model would be improved by explicitly considering individual exposed birds. If the performance metric were population decline, appropriate data could be developed to answer that question as well. All data should be carefully developed to match the scale of the assessment with the scale of the data.

ADDITIONAL PANEL COMMENTS

The Panel agrees with the registrant's statements (rationale) for focusing their probabilistic analysis on chemical exposure. We will never have the response data we truly desire (variability in non-target species sensitivity), but we can do a better job with respect to exposure. This information also lends itself nicely to exposure risk mitigation.

One difficulty with the registrant's exposure analysis is the use of time in cotton fields or margins as an indication of dietary intake within treated habitats. As was noted during the previous FIFRA SAP meeting on chlorfenapyr, the proportion of daily food requirements obtained within the treated habitats may be more a function of food availability than time spent foraging for it. For example, when food is abundant, a bird may spend only a short time in treated habitats meeting its daily food requirement which could, however, result in a potentially high chemical exposure, the opposite of what would be predicted based on time-in-field alone.

Another difficulty is the use of market share to reduce the potential number of treated fields. Whereas this may be appropriate for the regional or national assessment, it would not be at the local level (e.g., field, crop reporting district) where local use of the chemical may be much more extensive.

In terms of "natural" vs "unnatural" habitats, this is an inappropriate differentiation when it comes to the value of particular habitats to wildlife. This is because virtually all of our wildlife habitat has been altered by humans to some extent. With the loss of natural (optimum) habitat, many wildlife species have had to adapt to those that are less optimal and heavily manipulated by man. This does not mean that the latter are not important in wildlife conservation/protection. It means just the opposite: that we must look for ways of maintaining/managing theses habitats (e.g., agricultural lands) in a way that continues to provide habitat for species that utilize them

The probabilistic risk analysis simulates exposures for individual birds as a frequency distribution. The uncertainty about this distribution needs to be expressed. This is performed either by satisfying the sub-population or performing a 2-D Monte Carlo to separate uncertainty from variability.

The uncertainty in risk needs to be addressed. Currently, RQs greater than 1 cannot be interpreted. Suggestions are to either model the dose response, or plot the variability of exposure and compare against various toxicity bench marks. In the estimation of an RQ, how is background determined? The background is not zero if the birds are also exposed to residue from other pesticides and other toxics that have similar mechanisms of action.

LITERATURE CITED

Moore, DRJ. 1998. The ecological component of ecological risk assessment: Lessons from a field experiment. Human and Ecological Risk Assessment 4(5): 1103-1123.

O'Neill RV, DeAngelis DL, Waide JB and Allen TFH. 1986. A Hierarchical Concept of Ecosystems. Princeton University Press, Princeton, NJ.

Tacha, T. C., S. J. Schacht, R. R. George, and E. F. Hill. 1994. Anticholinesterase exposure of white-winged doves breeding in lower Rio Grande Valley, Texas. J. Wildl. Manage. 58:213-217.