UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, DC 20460

OFFICE OF CHEMICAL SAFETY AND POLLUTION PREVENTION

December 11, 2012

MEMORANDUM

SUBJECT: Transmittal of Meeting Minutes of the FIFRA Scientific Advisory Panel Meeting held September 11-14, 2012 on "Pollinator Risk Assessment Framework"

TO: Steven Bradbury, Ph.D.
Director
Office of Pesticide Programs

FROM: Fred Jenkins, Jr., Ph.D. Fred 12/11/12
Designated Federal Official
FIFRA Scientific Advisory Panel
Office of Science Coordination and Policy

THRU: Laura Bailey, MS Laura Bailey 12/13/12
Executive Secretary
FIFRA Scientific Advisory Panel
Office of Science Coordination and Policy

Frank Sanders
Director
Office of Science Coordination and Policy

Attached, please find the meeting minutes of the FIFRA Scientific Advisory Panel open meeting held in Arlington, VA on September 11-14 , 2012. This report addresses a set of scientific issues associated with "Pollinator Risk Assessment Framework."

Enclosure

cc:
Jim Jones
Louise Wise
Vicki Dellarco
William Jordan
Margie Fehrenbach
Keith Matthews
Donald Brady
Jack Housenger
Jennifer Mclain
Lois Rossi
Jay Ellenberger
Richard Keigwin
Mah Shamim
Reuben Baris
Joseph Decant
Kristina Garber
Thomas Moriarty
Keith Sappington
Thomas Steeger
Christina Wendell
Dale Kemery
Douglas Parsons
Vanessa Vu
OPP Docket

FIFRA Scientific Advisory Panel Members
Daniel Schlenk, Ph.D.
Kenneth Delclos, Ph.D.
Stephen Klaine, Ph.D.
Martha Sandy, Ph.D.
James McManaman, Ph.D.

FQPA Science Review Board Members

May Berenbaum, Ph.D.
Nina Fefferman, Ph.D.
Greg Hunt, Ph.D.
Rosalind James, Ph.D.
Nancy Ostiguy, Ph.D.
Jeff Pettis, Ph.D.
Mr. Jens Pistorius
Thomas Potter, Ph.D.
Arthur P. Schwab, Ph.D.
David Tarpy, Ph.D.

SAP Minutes No. 2012-06

A Set of Scientific Issues Being Considered by the Environmental Protection Agency Regarding Pollinator Risk Assessment Framework

September 11 – 14, 2012
FIFRA Scientific Advisory Panel Meeting
Held at the
Environmental Protection Agency Conference Center
Arlington, VA

NOTICE

These meeting minutes have been written as part of the activities of the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), Scientific Advisory Panel (SAP). The meeting minutes represent the views and recommendations of the FIFRA SAP, not the United States Environmental Protection Agency (Agency). The content of the meeting minutes does not represent information approved or disseminated by the Agency. The meeting minutes have not been reviewed for approval by the Agency and, hence, the contents of these meeting minutes 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 FIFRA SAP is a Federal advisory committee operating in accordance with the Federal Advisory Committee Act and established under the provisions of FIFRA as amended by the Food Quality Protection Act (FQPA) of 1996. The FIFRA SAP provides advice, information, and recommendations to the Agency 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 Environmental Protection Agency, Office of Pesticide Programs (OPP), and is structured to provide balanced expert assessment of pesticide and pesticide-related matters facing the Agency. FQPA Science Review Board members serve the FIFRA SAP on an ad hoc basis to assist in reviews conducted by the FIFRA SAP. Further information about FIFRA 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 Fred Jenkins, Jr., Ph.D., SAP Designated Federal Official, via e-mail at jenkins.fred@epa.gov.

In preparing these meeting minutes, the Panel carefully considered all information provided and presented by EPA, as well as information presented by public commenters.

TABLE OF CONTENTS

PARTICIPANTS 6

INTRODUCTION 9

PUBLIC COMMENTS.....10

SUMMARY OF PANEL DISCUSSION AND RECOMMENDATIONS.....11

DETAILED PANEL DELIBERATIONS AND RESPONSE TO CHARGE.....40

REFERENCES 98

SAP Minutes No. 2012-06

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

Pollinator Risk Assessment Framework

September 11-14, 2012
FIFRA Scientific Advisory Panel Meeting
Held at the
Environmental Protection Agency Conference Center
Arlington, VA

Daniel Schlenk, Ph.D.
FIFRA SAP Chair
FIFRA Scientific Advisory Panel
Date: DEC 10 2012

Fred Jenkins, Jr., Ph.D.
Designated Federal Official
FIFRA Scientific Advisory Panel
Date: DEC 10 2012

Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA)
Scientific Advisory Panel (SAP)
Pollinator Risk Assessment Framework
September 11 – 14, 2012

PARTICIPANTS

FIFRA SAP Chair

Daniel Schlenk, Ph.D.
Professor of Aquatic Ecotoxicology & Environmental Toxicology
Department of Environmental Sciences
University of California, Riverside
Riverside, CA

Designated Federal Official

Fred Jenkins, Jr., Ph.D.
US Environmental Protection Agency
Office of Science Coordination & Policy
FIFRA Scientific Advisory Panel
EPA East Building, MC 7201M
1200 Pennsylvania Avenue, NW
Washington, DC 20460
Phone (202) 564-3327
Fax (202) 564-8382
jenkins.fred@epa.gov

FIFRA Scientific Advisory Panel Members

Kenneth Delclos, **Ph.D.**
Pharmacologist
National Center for Toxicological Research
U.S. Food and Drug Administration
Jefferson, AR

Stephen Klaine, Ph.D.
Professor & Director
Environmental Toxicology & Biological Sciences
Clemson Institute of Environmental Toxicology
Pendleton, SC

Martha Sandy, Ph.D.
Senior Toxicologist & Chief
Cancer Toxicology & Epidemiology Section
Reproductive and Cancer Hazard Assessment Branch
Office of Environmental Health Hazard Assessment
California Environmental Protection Agency
Oakland, CA

James McManaman, Ph.D.
Professor & Chief
Section of Basic Reproductive Sciences
Department of Obstetrics & Gynecology, Physiology & Biophysics
University of Colorado, Denver
Aurora, CO

FQPA Science Review Board Members

May Berenbaum, Ph.D.
Professor & Department Head
Department of Entomology
University of Illinois at Urbana-Champaign
Urbana, IL

Nina Fefferman, Ph.D.
Associate Professor
Department of Ecology, Evolution & Natural Resources
The Center for Discrete Mathematics & Theoretical Computer Science
Rutgers University
New Brunswick, NJ

Greg Hunt, Ph.D.
Professor
Department of Entomology
Purdue University
West Lafayette, IN

Rosalind James, Ph.D.
Research Leader
Pollinating Insect-Biology, Management & Systematics Research Unit
US Department of Agriculture-Agricultural Research Service
Department of Biology
Utah State University
Logan, UT

Nancy Ostiguy, Ph.D.
Associate Professor
Department of Entomology
Pennsylvania State University
University Park, PA

Jeff Pettis, Ph.D.
Research Leader
US Department of Agriculture-Agricultural Research Service
Bee Research Laboratory
Beltsville, MD

Mr. Jens Pistorius
Scientist
National Examination Centre for Bee Poisoning Incidents
Julius Kühn InstituteFederal Research Centre for Cultivated Plants
Institute for Plant Protection in Field Crops and Grassland
Julius Kühn-Institut (JKI) [Institut für Pflanzenschutz in Ackerbau und Grünland
Braunschweig, Germany

Thomas Potter, Ph.D.
Research Chemist
US Department of Agriculture-Agricultural Research Service
Southeast Watershed Laboratory
Tifton, GA

Arthur P. Schwab, Ph.D.
Professor
Department of Soil & Crop Sciences
Texas A&M University
College Station, TX

David Tarpy, Ph.D.
Associate Professor & Extension Apiculturist
Department of Entomology
North Carolina State University
Raleigh, NC

INTRODUCTION

On September 11-14, 2012, the United States (US) Environmental Protection Agency (EPA) convened a public meeting of the FIFRA Scientific Advisory Panel (SAP) to address scientific issues associated with the Office of Pesticides Program's (OPP) proposed "Pollinator Risk Assessment Framework." Several sources have reported declines in certain pollinator species globally. A number of factors/agents have been hypothesized as potential contributors to recent declines in honey bee health in general. Currently, no factor has been identified as the single cause. Rather, the available science suggests that pollinator declines are a result of multiple factors which may be acting in various combinations. The purpose of this SAP was to advise OPP on the proposed Pollinator Risk Assessment Framework which focuses on a tiered process for quantitatively evaluating the potential risk to pollinators (using honey bees, *Apis mellifera,* as a surrogate) associated with the registered use of both systemic and non-systemic pesticides and the exposure and effects data needed to support that process. During this SAP, the US EPA provided an overview of the proposed process for quantifying potential risks of pesticides to honey bees. This overview reflected OPP's collective efforts with the Pest Management Regulatory Agency (PMRA), and the California Department of Pesticide Regulation (CalDPR).

Opening remarks at the meeting were provided by:
Steven Bradbury, Ph.D. (Director of the Office of Pesticide Programs)
EPA; Donald Brady, Ph.D. (Director of the Environmental Fate and Effects Division (EFED))

US EPA presentations were provided by the following staff:

Reuben Baris, MS
Joseph Decant, MS
Kristina Garber, MS
Thomas Moriarty, MS
Keith Sappington, MS
Thomas Steeger, Ph.D.
Christina Wendell, MS

Presentations from the Pest Management Regulatory Agency (PMRA), Health Canada and the California Department of Pesticide Regulation, California Environmental Protection Agency (CAL, EPA) were provided by:

Mary Mitchell (PMRA)
Richard Bireley (CAL EPA)

Public Comments

Public comments were provided by (listed in order of that they presented):

Dr. Ray McAllister, Dr. David Fischer, and Dr. Jay Overmyer on behalf of CropLife America
Mr. Pierre Petelle and Maria Trainer on behalf of CropLife Canada
Mr. William Kelly on behalf of the Center for Regulatory Effectiveness
Mr. Scott Schertz on behalf of the National Agricultural Aviation Association
Mr. Peter Jenkins on behalf of the Center for Food Safety
Ms. Nichelle Harriet on behalf of Beyond Pesticides
Mr. James Doan on behalf of Doan Family Farms
Dr. Michael Beavers on behalf of the California Agricultural Research Inc.
Mr. Rod Synder on behalf of the National Corn Growers Association
Mr. Steven McFadden of Independent Scientific Research Advocates on behalf of themselves

Summary of Panel Recommendations

1. **Section 2.2.1** of the white paper discusses the protection goals and associated assessment endpoints for assessing risks to honeybees (*Apis mellifera*). The protection goals are:
    - protection of pollination services;
    - protection of honey and hive product production; and,
    - protection of pollinator biodiversity.

   As described in the white paper, assessment endpoints are based on their ecological relevance, their susceptibility to known or potential stressors and their relevance to protection goals.

*a. Please comment on whether the assessment endpoints (e.g., population size and stability of managed bees, quantity and quality of hive products, and species richness and abundance) identified in **Table 1** in **Section 2.2.1** of the white paper are consistent with the Agency's protection goals. Please include a discussion of any additional assessment endpoints that may be necessary to meet those protection goals.*

Panel Summary

Overall the Panel believes that the Agency did an excellent job in writing the white paper. In regard to the assessment endpoints listed in Table 1 of the white paper, the Agency should clarify the terminology in this Table. For example, "managed bees" should be clearly defined as commercial bees including bumble bees, alfalfa leafcutting bees, mason bees, and any other bee species that may become managed in the future. Also the term population size is not explicitly characterized in Table 1. Thus, the Panel recommends that this term be specifically defined as "the numbers of colonies."

The Panel also believes that the Agency's protection goal of "contribution to pollinator biodiversity" is too broad. This is because pollinators are comprised of very large numbers of organisms including many bee species and other pollinating insects (*i.e.* some butterflies, pollen wasps, pollen beetles, *etc*.) and animals (*i.e.* some birds, and bats). Considering the large number of different pollinators and taking into account the focus of the Agency's white paper, the Panel advises that the Agency revise this protection goal to be "contribution to bee diversity."

*b. Please comment on whether the measurement endpoints at the level of the colony (e.g., colony strength and survival, contamination of pollen and nectar and species richness and abundance) identified in **Table 1** are consistent with the assessment endpoints identified in the table and any additional assessment endpoints discussed in Part "a" of this question. Please include a discussion of any additional measurement endpoints that may be necessary to represent those assessment endpoints.*

Panel Summary

Regarding the measurement endpoint "colony strength", the Agency should take note of the fact that 98% of native bee species are solitary, and thus a colony measurement endpoint does not include solitary bees. The Panel believes that this endpoint is acceptable for use as long as this caveat is considered. The Panel advises the Agency to add "colony development" as a measurement endpoint along with colony strength and survival. The Panel also recommends that the Agency explicitly identify pesticide residues in honey, pollen, wax, propolis, and royal jelly as a measurement endpoint.

The assessment endpoint of "species richness and abundance" is an appropriate goal for assessment; however, the measurements associated with this endpoint are not described in the white paper. Moreover, the Panel notes that there is no means of assessing species diversity using only one surrogate species, the honeybee. It is also important to note that the honey bee is a domesticated organism that is not native to the Americas. Furthermore, a substantial amount of research indicates that the presence of honeybees can have a harmful impact on native species of bees (Badano, 2011, Dohzono, 2010, Thomson, 2006 & 2004, Roubik, 2001, & Schaffer, 1983).

The Panel also recognizes that "within-species richness and diversity" is a good assessment endpoint to address. Thus, the Panel suggests that the Agency explore measurement of genetic variation within the species as an approach for assessing species richness and diversity. Genetic population modeling reveals genetic variation within a species and can be a valuable means of determining long term effects to population size.

*c. Please comment on whether the measurement endpoints at the level of the individual bee (e.g., individual adult and larval [brood] survival, queen fecundity, brood emergence success, worker longevity) identified in **Table 1** are consistent with assessment endpoints identified in the table and any additional assessment endpoints discussed in Part "a" of this question. Please include a discussion of any additional measurement endpoints that may be necessary to represent those assessment endpoints.*

Panel Summary

The Panel agrees that individual worker survival is a useful individual measurement. In addition, the Panel recommends that the Agency conduct larval survival assays. These assays will serve an important role in ascertaining more subtle effects of pesticides on bees. The Panel also recommends for the sake of clarity that the term "brood success" in Table 1 be replaced with "larval survival, delayed development and pupal survival." An additional term that should be

renamed to provide more clarity is brood size. Brood size should be changed to brood "nest" size (defined as the area of the sealed brood in a colony). "Queen fecundity" (the total number eggs laid in a defined time period) is another useful endpoint, but it is difficult to measure. Queen reproductive success (the size and quality of the brood nest, or survival of eggs through the larval, pupae and adult stage) is a superior endpoint. In regard to "worker bee longevity" as a measurement endpoint, the Panel notes that it is difficult and unreliable to measure and therefore, not recommended.

2. **Section 2.2** of the white paper discusses a series of conceptual models for assessing risks of honeybees resulting from pesticide applications. These models depict the nature of the stressor (*i.e.*, nonsystemic and systemic pesticides that are applied as foliar treatments, soil and/or seed treatments, or trunk injection), its source (*e.g.*, direct deposition on bees or pollen), the exposure media (*e.g.*, residues on plants, residues in/on pollen/nectar), receptors (*e.g.*, foraging bees, developing brood) and attribute changes (*e.g.*, reduced [bee] population, reductions in the quantity/quality of honey).

*a. Please comment on whether the conceptual models depicted in **Figures 4** through **8** are consistent with the protection goals and assessment endpoints identified in **Table 1** and discussed in Question 1.*

Panel Summary

The Panel believes that the Agency's conceptual models effectively assess pollinator risk while providing considerable detail. However, the Panel recommends some revisions. Examples include the potential receptor pathway/exposure route via wax and propolis specifically during the larval stage and the potential exposure due to consumption or contact with pesticide contaminated water, soil, and dust. Other suggested revisions are identified in Figures 4-9 of the "Detailed Panel Deliberations and Response to Charge" section of this report.

3. The focus of the white paper and proposed risk assessment process on the honeybee reflects two important factors: 1) honeybees are considered one of the most important pollinators globally from both a commercial and ecological perspective; and 2) standardized test methods for evaluating exposure and effects of chemicals in a regulatory context are more developed with the honey bee compared to non-*Apis* bees. Both the Introduction (**Section 1**) and the Problem Formulation sections **(Section 2**) of the white paper indicate that honey bees have historically served as a surrogate for other beneficial insects. As discussed in **Section 5.3** of the white paper, there is uncertainty regarding the extent to which honey bees serve as surrogates for native species, especially where life history strategies and differential sensitivity across species render native species more or less vulnerable to pesticides.

a. Please comment on the extent to which the assessment of risks to the honey bee may serve to meet the protection goals identified in the white paper (i.e., protection of pollination services; protection of honey and hive product production; and protection of pollinator biodiversity).

b. Until guidelines are developed for testing non-Apis species of bees, please comment on the extent to which the honey bee may or may not serve as a reasonable surrogate for non-Apis bees, given the differences in life history strategies and potential different sensitivities to pesticide toxicants. Please include a discussion of which types of non-Apis bees may be particularly well represented by either the individual-level or colony level endpoints identified in Table 1 of the white paper; as well as which types of non-Apis bees may not be as well represented, and therefore, may be the focus of potential areas for future research.

Panel Summary

Subparts Questions 3a. and 3b. are closely connected; therefore, the Panel prefers a joint response to them. The Panel believes that honey bees are a reasonable surrogate. They are easier to obtain than most other bees, and testing methods are more readily available. However, if honey bees are to serve as a surrogate species, the Agency must recognize that this species is merely a representative of other bee species. It is important to consider the differences between honey (*Apis mellifera*) and other bees. For example, some non-Apis bees may consume proportionately more pollen and less nectar. If pesticide residues in pollen are dissimilar from nectar residues these differences in consumption rates may impact dietary exposure estimates. Also, honey bees do not frequent soil while other bees do. For example, many solitary bees build nests with soil; thus they may be exposed to pesticide residues in soil. The Agency's white paper does not have risk diagrams that take into account this potential exposure pathway when non-systemic pesticides are applied to the soil (See Panel Proposed Figure 8 in Question 2 Part a. of the Detailed Panel Deliberations and Response to Charge Section). In addition contact toxicity assessments for other soil applied pesticides are lacking. An additional consideration that the Agency should address is that honey bees have the capability of recruiting more workers from the young nurse bees, if necessary. In the case of solitary bees, all females are queens. Thus, adult mortality could have a greater impact on these bee populations in the following year than would be the case for honey bees.

The Agency may want to consider testing which entails other bee species. The commercially available alfalfa leafcutting bee, *Megachile rotundata* and *Osmia* spp. (mason and orchard bees)

would likely be relatively easy to include in Tier I testing. Bumble bees (also commercially available) (*Bumbus* spp.) may be appropriate for use in Tier II tests although it would be difficult to assess larval development and brood size or larval mortality. The Panel recommends that the Agency require testing on at least one additional species to address the stated goal of protecting diversity.

4. **Contact Exposure.** The exposure characterization of the white paper (specifically, **Section 3.1.1.2**) proposes a screening-level (Tier I) approach for quantifying contact exposure to foraging bees for pesticides applied via foliar spray. This proposed method is based on the maximum of residue values on honey bees from a field study conducted by Koch and Weisser (1997). The white paper also discusses a method based on the T-REX (Terrestrial Residue Exposure model) upper-bound concentration for arthropods directly sprayed with pesticides while located directly on a treated field. Although the second method is not proposed for the Tier I exposure assessment for honey bees, it could be used to assess contact exposures to other insect pollinators.

a. *Please comment on the strengths and limitations of the proposed approach for assessing contact exposures to honey bees in Tier I exposure assessments (i.e., 2.7 µg a.i./bee per 1 lb a.i./A), which is based on the honey bee specific maximum concentration reported by Koch and Weisser (1997).*

Panel Summary

The Panel agrees that contact exposure is likely a significant pesticide exposure pathway. The approach described in the white paper for determining contact toxicity has many merits. It also has several limitations. Given the general lack of actual contact exposure data, a model appears necessary to estimate contact exposures. An example of the strength of the modeling approach described is that the concentration reported was the highest value (normalized). Consequently, this value should be protective. Major weaknesses of this approach are that the primary study is 15 years old and investigators did not use a formulated pesticide to evaluate exposure. Follow-up studies may be needed to assess study relevance and validity.

b. *Please comment on the potential utility of the T-REX upper-bound residue value (i.e., 12 µg a.i./bee per 1 lb a.i./A), for a broader number of arthropod species to represent contact exposures to honey bees and to other insect pollinators that are directly sprayed with pesticides.*

Panel Summary

A primary strength of T-REX is that the data that the model uses were derived from pesticide residue measurements on arthropods thus they are expected to be similar to what might be expected for both honey and non-*Apis* bees. A main weakness is that T-REX was developed using data from only carbamate and organophosphate insecticides. Other pesticides, *e.g*. the neonicotinoids, have different chemical properties and are applied in different ways; thus they may respond differently. Studies are needed to assess the potential differences in the magnitude of residues as a function of pesticide and formulation properties and mode of application.

Overall, considering the strengths and weakness of the T-REX approach, the Panel supports use of this model as a means of estimating the interception of a pesticide during a spray application to a crop. One panel member also recommends that the Agency considers a more direct approach to assessing potential exposure based on the surface area of adult honey bees. Values in the 2.4 cm^{-2} per bee range have been reported in the literature (Roberts & Harrison, 1999). Exposure estimates can be derived by multiplying the surface area times the spray rate. Some consideration for the rate of spray interception may also be included in the calculation. Using the default value of 100% interception will result in exposure estimates with a factor of 2X of the exposure estimates derived with the TREX model.

5. **Dietary Exposure (Consumption Rates).** As discussed in the effects assessment section of the white paper (**Section 4.1**), acute oral toxicity data (LD_{50}) are necessary for adult and larval bees in order to characterize the potential risks of a pesticide. Because these toxicity data are expressed on a dose basis (*i.e.*, µg a.i./bee), it is necessary to convert estimated concentrations of pesticides in food (expressed as mg a.i./kg) into doses. Honey bees fulfill their nutritional requirements through consumption of nectar, honey, and bee bread (pollen/honey). In addition to requiring bee bread and nectar or honey, bees also require royal jelly and brood food (jelly) to fulfill their nutritional requirements. In the proposed approach, it is assumed that exposures through consumption of nectar and pollen are conservative representations of potential exposures through consumption of honey and bee bread, respectively. This approach is likely to be conservative because it assumes that pesticides do not degrade while honey and bee bread are stored in the hive and have undergone some degree of processing (*e.g.*, fermentation). For bees that consume honey, it is assumed that the estimated pesticide exposures can be related back to the original concentration in nectar by accounting for the amount of sugar consumed by bees. It is also assumed that pesticide exposures through consumption of pollen and nectar are protective of pesticide through consumption of royal jelly and brood food given that empirical data indicate that pesticide concentrations in royal jelly are >100 times lower than concentrations in food consumed by nurse bees. In the proposed approach, pesticide doses received by bees are calculated using nectar and pollen consumption rates for larval and adult worker bees. As discussed in detail in **Appendix 1** of the white paper, the proposed values for larvae and adult workers are based on an analysis conducted by EPA, which built upon work published by Rortais *et al.* (2005), Crailshaim *et al*. (1992 and 1993) and others. For larvae, the proposed total food consumption rate is 120 mg/day, which is based on the total daily consumption of pollen and nectar (based on honey consumption) by larvae during day 5 of the uncapped larval life stage. For adult worker bees, the proposed food consumption rate is 292 mg/day, based on food consumption rates of nectar foraging bees, which are expected to receive the greatest dietary exposures among different types of worker bees. In addition, as discussed in **Appendix 1**, it is likely that these food consumption rates are protective of drones and queens.

a. *Although bee larvae typically consume processed foods in the form of royal jelly and brood food throughout much of their development, they also consume honey and pollen during the last two days of the uncapped period. Please comment on the proposed use of nectar and pollen consumption rates of larvae during the last day of the larval developmental stage. Please include a discussion of the conservatism, strengths and limitations of this approach as well as a discussion of how this value may or may not correspond to data generated from larval toxicity endpoints.*

Panel Summary

The Panel generally concurs that unprocessed pollen and nectar consumption represents a worst case scenario as well as a conservative estimate when compared to processed foods such as brood food and royal jelly. Regarding larval assays, as demonstrated by published literature larval tests can be done using dilute royal jelly as food. Further, spiking royal jelly with pesticides will allow the larvae to be orally exposed. Although the Agency's proposed method appears conservative the Panel questions the Tier I laboratory larval feeding concentrations. As identified in the white paper the Agency assumes that pesticide residue concentration in royal jelly is 100-fold lower than in the pollen and nectar that a bee ingests to prepare royal jelly. The Panel recommends using the original estimated adult ingestion concentration because the 100-fold lower value used for royal jelly was derived from a small number of measurements. There is relatively high uncertainty in concentration estimates associated with this small sample size. Concerning larval food intake, the Agency assumes an intake value of 120 mg per day. This is likely a highly conservative value. As indicated in published studies, larvae of the first up to the last larval stages would consume less food. The Panel also notes regarding food consumption that the total food consumption during the larval stage may be the most appropriate means of estimating exposure. This approach would likely provide a more accurate pesticide exposure estimate of both honey and non-*Apis* bee larvae. The Panel notes that honey and solitary bee larvae have 7 and 1 day time periods respectively over which food is collected for larval consumption.

b. *Please comment on the strengths and limitations of basing the Tier I screen for adult honey bees on food consumption rates of nectar foraging bees, including a discussion of the conservatism of this approach, and how it relates to other types of worker bees and castes*.

Panel Summary

Panel members believe that Tier 1 screening using nectar consumption only is generally appropriate. However, this approach may not be for newly-emerged bees since these young bees may consume very large amounts of pollen or nectar. Since Tier I studies generally use random aged bees exposure estimates based on nectar consumption appears appropriate. The Panel notes that the proposed value for nectar consumption, 292 mg per day will likely be highly protective for nectar consumption of non-Apis species.

The Panel notes that since queens and drones do not normally feed themselves and are fed by worker bees, the proposed Tier I food consumption values are likely acceptable for determining LD_{50} values. However, the Panel advises the Agency to also consider the high daily feeding rate

of the queens. This may result in much higher exposure when compared to the workers alone. To some degree this is moderated by lower pesticide concentration levels in the brood food fed to the queens. With regard to adjusting for castes and species differences, the Panel recommends computing exposure as a function of bee weight as opposed to a per bee basis, as is done in the white paper.

c. Please comment on the assumption that exposures through consumption of nectar and pollen are conservative representations of potential exposures through consumption of honey, bee bread, brood food and royal jelly, all of which represent processed foods.

Panel Summary

The Panel agrees that the proposed nectar and pollen consumption rates provide conservative representations of exposure levels compared to processed brood food (honey and bee bread) and royal jelly. However, the Panel notes that this conclusion is not easily extrapolated to non-*Apis* native bees because their larvae receive provisions of pollen and relatively little nectar. Additionally, native bee larvae are not fed brood food or royal jelly.

6 Dietary Exposure. The dietary exposure methods described in **Section 3.1** of the white paper differ in the nature of the estimated concentrations in pollen and nectar consumed by bees. For foliar spray applications, the proposed approach involves the use of the tall-grass residue value from the T-REX model (v. 1.5) as a surrogate for pesticide concentrations in nectar and pollen. For soil treatments, the white paper proposes the use of the Briggs' model, which is designed to estimate pesticide concentrations in plant shoots resulting from plant uptake of pesticides from treated soil. Estimated pesticide concentrations in plant shoots are proposed as a surrogate for concentrations in pollen and nectar. For seed treatments, the white paper proposes the use of the International Commission for Plant-Bee Relationships' (ICP-BR) 1 mg a.i./kg concentration as an upper-bound concentration in nectar and pollen. The paper explores the strengths and limitations of each method relative to the ability to derive reasonably conservative estimates of pesticide exposures to bees, with a focus on how well the estimates relate to empirically based measures of pesticides in pollen and nectar from crops treated with pesticides.

a. Foliar spray: Please comment on the analysis presented in Section 3.1.1.1, with a focus on the extent to which the T-REX tall-grasstall-grass upper-bound residue may serve as an adequate surrogate to represent upper-bound pesticide concentrations in pollen and nectar of flowers that are directly sprayed with pesticides.

Panel Summary

The Panel believes that due to the limited amount of data on pollen and nectar pesticide residues, the T-REX or other models are needed to predict upper-bound pesticide concentrations in these exposure media. The Panel strongly recommends a focused effort to collect more residue data on pollen and nectar to determine whether the T-REX estimates are an appropriate surrogate for these materials. The Panel notes several strengths and weaknesses associated with the use of T-REX. Examples of these strengths and weaknesses include:

Strengths

- T-REX is a model currently endorsed and used by US EPA in Tier I studies.
  The model was developed using experimental (empirical) observations on stems of plants in 11 direct spray studies.

Weaknesses

- Residue data on nectar and pollen data are limited. Thus, there is difficulty and high uncertainty associated with determining whether T-REX serves as an appropriate surrogate for assessing residue values.
- Linear extrapolation to 1 lb a.i./acre has not been validated. Changes may be non-linear.
- Crop specific factors are not accounted for including differences in plant surface composition and characteristics of grass stems (T-REX model) that impact pesticide retention and other plant parts (e.g. blossoms).

b. Soil applications: Please comment on the analysis presented in Section 3.1.2, with a focus on the extent to which the Briggs' model may generate estimates of pesticide exposure in plant stems that can represent upper bound pesticide concentrations in pollen and nectar of flowers.

Panel Summary

The Briggs' model has several strengths, uncertainties, and weaknesses. A few examples of these include:

Briggs' Model Strengths:

- The Briggs, among available plant uptake equations, was determined by the United Kingdom's Environmental Agency to represent the best fit to empirical data.
- This model is used in other Agency models such as PRZM.
- The range in log K_{ow} values, -0.57 to 4.6, used to derive the Briggs' equation includes a large number of pesticides including the neonicotinoids.

Briggs' Model Weaknesses:

- The model examines nonionic compounds only. Pesticides that ionize will not be well represented by this approach.
- Pesticides tested were limited to carbamates and phenyl-ureas and only one plant species (barley). It is unknown whether the model holds for other pesticides or plants.
- Soil parameters used in the Ryan modification may not be known, particularly soil moisture content. The fact that other parameters like bulk density and f_{oc} are spatially dependent will also add uncertainty to results.

Burken and Schnoor (1998) assessed the Briggs' model using another plant species and a much broader array of chemicals. The Panel recommends the Burken and Schnoor (1998) analysis of the Briggs' model be incorporated into the Agency's evaluation.

c. *Soil applications: Please discuss the relative strengths and limitations of the 1 mg a.i./kg value and the soil uptake model (the Briggs' model) proposed in the white paper as Tier I screens, including consideration of the extent to which this method may generate conservative Tier I estimates of dietary exposures to bees. Does the Panel conclude that the one approach may be better suited to specific types of assessment scenarios? If so, please elaborate. Alternatively, if both approaches are equally suited for a Tier I screen, please provide guidance on how to be capture variability and uncertainty in the exposure estimates using the two approaches.*

Panel Summary

The Panel concludes that although the Briggs' approach has multiple uncertainties and weaknesses, it is a more rational and defendable approach to estimate Tier I pesticide dietary exposures to bees than the European and Mediterranean Plant Protection Organization (EPPO) default value (*1 mg a.i./kg).* This value does not appear to take into account pesticide properties as does the Briggs' method. Consequently, the Panel believes that the Briggs' approach is stronger mechanistically, and is a superior method for estimating pollen and nectar pesticide concentrations.

d. *Seed Treatments: Please comment on the analysis presented in Section 3.1.3, including a discussion of the strengths and limitations of the use of 1 mg a.i./kg value as an upper-bound concentration for pollen and nectar of seed-treated crops.*

Panel Summary

There are a few strengths, but many more weaknesses associated with use of the 1 mg a.i./kg value as an upper-bound concentration for pollen and nectar of seed-treated crops. The Panel concludes that the value is likely conservative, but believes that a comprehensive assessment of measured data is needed before this approach can be confidently used. The Panel recommends that a concerted effort be made to compile all available data to assess potential ranges in observed pesticide concentrations in pollen and nectar with different plants and pesticides.

e. *Please comment on other approaches or data that should be considered for estimating upper-bound estimates of pesticide residues in pollen and nectar as a Tier I screening-level assessment for pesticides applied via foliar spray, soil application or seed treatment.*

Panel Summary

As indicated in the Agency's white paper, there are other methods for estimating Tier I upper-bound pesticide residue estimates in pollen and nectar resulting from foliar spray, soil application, and seed treatments. They are for example:

- Direct measurement of the pesticide in question on the plant parts of interest (nectar, pollen, etc.).
- Direct measurement of the pesticide in question in other plant tissues (stems, leaves, etc.).
- Direct measurement of surrogate pesticides in plant tissues.

As noted previously, (see response to Question 6. d.) the Panel strongly advises a data call-in to assist in evaluation of these approaches and determining appropriate upper-bound values for use in Tier I assessments.

7. **Consideration of other Exposure Pathways**. The proposed measures of exposure are based on what are believed to be the primary routes, *i.e.*, direct contact and ingestion of contaminated pollen and nectar. Additional routes of exposure are considered (*e.g*., dust, drinking water), but not included in the proposed Tier I exposure assessment method. As discussed in **Section 3.1.4.1** of the white paper, effective quantitative screening methods for estimating exposures through contaminated dust are not discernible at this time. The most effective management of bees' exposure to pesticides through dust appears to be through pesticide application (*e.g.*, stickers) and seed planting practices, especially since dust exposure is expected to be a concern for only a limited number of pesticides and application scenarios. In regard to pesticide exposures through drinking water, **Appendix 2** presents an analysis of potential exposures to bees through various sources of water to support the exclusion of drinking water exposure in the Tier I screen. The results of this analysis indicate that if bees consume the majority of their water from puddles or ponds, the exposures relative to dietary and direct spray are insignificant. The preliminary analysis indicates that if bees drink a substantial amount of water from guttation fluid or dew, conservative exposures may be similar to or even exceed pesticide exposures through the diet or direct spray. Further investigation concluded that pesticide exposures through dew and guttation fluid are not expected to be as significant when compared to diet, primarily because they are not likely to consistently drink a substantial amount of water from these sources.

a. Please comment on the strengths and limitations of basing the Tier I exposure method on contact and diet. Does the Panel agree that for the majority of pesticide applications, the primary exposure routes for bees will be represented by contact and diet?

Panel Summary

In general the Panel concurs that diet and direct contact are the primary pesticide exposure pathways. However, the Panel considers exposure from seed treatments to also be a potential exposure pathway particularly since bee mortalities during planting of seeds treated with certain insecticides are widely reported. Exposure in this case is expected to occur via direct contact and

oral uptake with dust generated during the planting of pesticide treated seed, granular formulation application. The Panel also believes that pesticide exposure via contact and consumption of pesticide contaminated water should be included in Tier I assessments.

b. Dust: If the Panel believes that this exposure route should be quantitatively included in the Tier I exposure method, for the relevant application type(s) (i.e., seed treatment), please discuss the data that may be needed to address the exposure route quantitatively.

Panel Summary

The Panel concurs with the white paper findings that bees may be exposed via contact to pesticides in dust while in flight or when visiting flowers and other vegetated surfaces where dusts have deposited. The Panel also recognizes that there have been widespread reports of honey bee mortality associated with dust emissions during planting of pesticide treated seed. Thus, the Panel notes that in many cases deaths due to pesticide exposure were confirmed by detection of pesticide concentrations above lethal limits in bees.

The Panel acknowledges that no matter what the dust source, an exposure assessment is challenging since there are many variables involved and little published data. Recently published work may provide some insight with regard to pesticide emissions associated with abraded seed coatings. For example, Tapparo et al. (2012) reported on the percentage of applied pesticide that may be emitted in planter exhaust from corn seed coated with various commercial insecticide treatments. Such data are needed to evaluate exposure estimates. One panel member notes that consideration of dust exposure and effects as well as potential mitigation measures should be addressed by higher tier testing and/ or risk management measures.

c. Drinking Water: Please comment on the analyses, discussions provided in Appendix 2 of the white paper and the conclusion that pesticide exposure to bees through drinking contaminated water is not expected to be a major route of exposure when compared to contact (following foliar spray applications) and diet. If the Panel believes that this exposure route should be quantitatively included in the Tier I exposure method, for the relevant application type(s) (i.e., foliar spray, soil treatment, seed treatment, or trunk spray), please discuss why and what data may be needed to address the exposure route quantitatively.

Panel Summary

Honey and other bees may be exposed to pesticides in water contaminated with pesticides via direct water consumption and/or during contact when worker bees carry water in their "honey stomachs" to the hive for use in hive evaporative cooling. The Panel notes that water exposure assessments in the white paper are limited to estimates of drinking water consumption. The Panel concurs that potential exposures should also include other types of water contact. However, the Panel believes that even when both consumptive and non-consumptive water contact are assessed, it is assumed that worker bees absorb 100 % of the pesticide transported in water in honey sacks. Potential exposures to water from contaminated ponds appear small when compared to dietary or contact exposures described for Tier I. In regard to the potential use of puddles in farm fields as bee water sources, the white paper also indicates potential exposures

were likely very small when compared to diet and direct contact exposure in Tier I estimates. Regarding this point, the Panel disagrees and notes that values with incorrect units were used in Equation 2-1 found at the top of page 210 in Appendix 2 of the white paper. The Panel believes that the equation is written correctly, but careful attention must be paid to input parameter units. Failure to do so in the computations described in the white paper resulted in 1000-fold underestimates of pesticide concentrations in puddle scenarios. Use of appropriate input values indicates that pesticide exposure via consumption or contact with water in puddles on or adjacent to pesticide treated fields may be relatively high. These estimates appear to be within the same range as reported for both dietary and contact exposure. The Panel also proposes an alternate approach for estimating pesticide concentrations in field puddles that was develop for the Agency's "ECOFRAM (Ecological Risk Assessment Framework; US EPA, 2012). Use of this model may provide further insight into potential pesticide exposures via this pathway.

Concerning guttation water, the Panel believes that guttation water may not be an important water source for bees since the drops are usually found in the morning only. However, the Panel notes that there is uncertainty regarding the extent to which honey bees may be exposed to guttation water and cautions that published studies show that residue levels in guttation water may be high.

d. Other Routes: Please identify and discuss additional exposure routes (if there are any besides contact with dust and consumption of drinking water) that would contribute significantly to pesticide exposure of bees and explain how and why such exposures could be considered quantitatively in establishing the Tier I exposure value.

Panel Summary

The Panel believes that a potentially significant route of exposure for some non-*Apis* species is direct contact with pesticide contaminated soil. Many solitary bee species nest in the ground. Also some cavity nesting bees use mud to wall off and divide the cells in their nest. Since the Agency currently does not have models for estimating these types of pesticide exposures the Panel recommends that research be conducted to directly address these exposures.

8. **Tier I Effects Assessments.** As discussed in the Problem Formulation (**Section 2.2.1**), the assessment endpoints for the ecological risk assessment of bees involve maintaining honey bee population size, stability of managed bees, quality and quantity of hive products, species richness and abundance. In order to use the results of toxicity studies quantitatively in risk assessment, it is important to identify specific endpoints which will be measured in toxicity tests as these measurement endpoints must have clear linkages to assessment endpoints. As indicated in **Table 1** of **Section 2.2.1**, at the individual bee level (which is the focus of the Tier I assessment), measurement endpoints relevant to these assessment endpoints include: individual survival, adult bee longevity, brood size, brood success, and queen fecundity. The acute and chronic toxicity tests with larvae and adults can be used to quantify effects of pesticides on all of these endpoints, with the exception of queen fecundity (which would require an egg laying study involving the queen). The focus of the chronic toxicity tests with larvae and adults is on mortality that may occur during the tests. Potential impacts of a pesticide on brood size and success can be assessed by determining whether there is decrease in the number of brood (*i.e.*, larvae) following a chronic exposure of larvae to that pesticide. Potential impacts of a pesticide on adult survival and longevity can be assessed by determining the mortality and the decrease in the life spans of adult bees following chronic exposures to the pesticide. The notable limitation to the proposed chronic toxicity endpoints is that they do not include measures of queen fecundity.

a. *Please comment on the extent to which currently available bee toxicity tests, which focus primarily on mortality/survival, serve as an effective Tier 1 screen.*

Panel Summary

The Panel notes that the Agency provided a clearly articulated, detailed, and thorough description of its Tier I bioassays in the white paper. In regard to suggested enhancements to the Tier I assays, the Panel provides the following considerations. Firstly, the Panel believes that these bioassays do not account for the subtle variation among individual bees. Thus, the Panel suggests that standardizing to honey bee weight seems to be a useful approach that would translate more effectively to other bee castes and especially other bee species. Secondly, the Panel notes that the current Tier I screens (which measure the percent mortality after 48 to 96 hrs) do not evaluate the temporal variation on final mortality. Thus, the Panel recommends that standard survival analyses would provide more information, as it would account for not only the percent mortality at each time point measured, but it would also statistically compare the slopes of the mortality curves. This information could be more readily applied in the social context of a bee colony. Thirdly, the Panel suggests that the Agency consider employing two types of Tier I honey bee cage studies. Each would be informative in different ways. These two include a one-time application study and a continuous exposure by feeding study. The Panel also advises the Agency to consider the physiological differences between summer and winter honey bees, and short-duration colony events (e.g. increased honey-store consumption during smoke application), and how the timing of pesticide applications could result in different outcomes due to these differences. Lastly, the Panel recommends that the Agency recognizes the genetic differences among populations as well as the need to consider the behavioral (sub-lethal) abnormalities that may occur from pesticide exposures. Since no trigger values are established for sub-lethal

effects, the risk assessor should have the flexibility to ask for additional data or higher tier tests when concern is raised on the basis of sub-lethal effects observed in laboratory studies.

b. Please comment on additional measurement endpoints (e.g., growth) which should be considered in future modifications of Tier I test protocols and which are appropriately linked to the proposed assessment endpoints. Given that the queen is the reproductive unit of the colony, please comment on methods to evaluate effects on individual queens, considering practical limitations of testing with queens.

Panel Summary

Currently there are no methods available for assessing effects on individual queens. However, the Panel notes that quantifying reproductive potential in queens can be done in many different ways including evaluating ovariole number, morphometrics, and pheromone production (c.f., Delaney et al. 2011).

In regard to additional measurement endpoints to be considered for Tier I protocols, the Panel suggests the following:

- A temporal component to Tier I bioassays, using survival statistics to quantify slopes as well as final percent mortality (see Question 8a).
- Quantification of immuno-competence on adults (and larvae) to assess the possible long-term sub-lethal consequences of exposure (however, bearing in mind the caveat that little is known about the linkages of the immune system to colony health at this point to make such inferences very strong).
- Measurement of motility, social interactions, and other behavioral changes (in addition to the survival measurement currently assessed by the Agency).Addition of a Proboscis Extension Reflex (PER) and learning bioassay. PER is a well vetted and commonly performed bioassay for learning and memory, which uses the instinctive reflex of a harnessed adult worker honey bee sticking out her proboscis when her antenna contacts a droplet of sucrose solution. However, one panel member disagrees that PER tests are currently very useful in risk assessments as they assert that there is a lack of available data which link these tests to the realistic situations that bee colonies encounter.
- Investigation into the development of a study that would test the impact of pesticides on pheromone production.
- Evaluation of viability of sperm within queens.
- Analysis of queens' fecal materials (a test that can provide some useful information and which can be done without sacrificing the queens).

9. Tier I Larval Toxicity Testing. Section 4.1 of the white paper discusses new data requirements for the screening-level effects assessment and recommends obtaining and using larval toxicity data on individual bees. The paper specifically identifies the assay initially proposed by Aupinel *et al*. (2007) as one methodology for quantifying acute oral larval toxicity in the Tier I screen. These assays rely on feeding bees a sugar solution which has been spiked with the test material; however, this *in vitro* method of feeding larvae differs from the process by which the larvae would typically be fed within the colony environment, *i.e*. by nurse bees secreting either brood food or royal jelly.

a. Please comment on the extent to which the Aupinel et al. (2007) in-vitro method serves as an appropriately conservative estimate of Tier 1 acute oral exposure of honey bee larvae to pesticides, given differences in this test design from actual in-hive exposure conditions (e.g., during the first 3 days of the larval development stage larvae consume royal jelly and brood food) and the uncertainty regarding the extent to which larvae rely exclusively on pollen/nectar as opposed to royal jelly/brood food.

Panel Summary

As mentioned in the Agency white paper, larval toxicity *in vitro* testing is complicated by factors such as genetic differences among colonies and difficulties synchronizing larval age. An additional complicating factor is the uncertainty regarding how larval food (*i.e.*, royal jelly/brood food, honey and beebread) may alter the responses of pesticides.

An additional concern is that an acute dose for a larva may be biologically difficult to interpret. This is because larvae receive food on a continual/chronic basis. Thus, the Panel suggests that with *in vitro* larval tests it is preferred that larvae are given contaminated food every day on a chronic basis as opposed to being fed on an acute basis in which the larvae are only given contaminated food once on Day 4, and on the other days uncontaminated food.

b. Please comment on the extent to which pesticides would be more or less bioavailable using the synthetic matrix relied on for feeding developing bees in this in vitro method.

Panel Summary

Bioavailability is directly impacted by the characteristics of the pesticide, larval metabolism, and food composition. Currently, there is an inadequate amount of data for the Panel to assess this question. For example, the Aupine *et al*. (2005) acute larval toxicity test used an organophosphate, dimethoate (which has relatively high water solubility). It is not clear how other types of pesticides will behave in the synthetic matrix.

c. Please comment on the extent to which the absence of trophallaxis (i.e., the transfer of food/fluids between colony members) may render larvae more or less vulnerable to pesticides.

Panel Summary

The Panel deems the impact of trophallaxis on pesticide resistance or susceptibility as completely unknown. To the Panel's knowledge, the understanding of this impact represents an absolute data gap.

d. Please comment on alternative methods for testing individual larvae that may be appropriate for quantitative use in a Tier I screening-level assessment.

Panel Summary

Published studies on alternative methods for quantitatively testing larvae are rare. Even rarer are literature reports of alternative methods for testing larvae quantitatively. One approach described by Atkins and Kallum (1986) involved treating larvae in a colony by micropipetting pesticides directly into the cells.

There has not been a common use of pupal bioassays for pesticide toxicity testing. However these bioassays have been commonly used in other milieux (e.g. pathogenicity of fungi.) Such studies may be reproducible and provide some informative means of developing pupal bioassays testing pesticides.

There are available data which describe the larval transcriptomes and methylomes of honey bees (Johnson *et al*. 2009). These data could provide a useful means of developing quantitative measures of effects of toxins on larval development that might be useful as alternative indicators of larval actions/health.

e. Typically acute toxicity tests are concluded between 48 – 96 hrs. Please comment on the appropriate duration of toxicity tests for assessing acute toxicity to individual larval and adult bees.

Panel Summary

A two to four day time period toxicity test presents some problems for adequately interpreting larval survival. This is because the four-day endpoint represents only 80% of the period of active feeding. Thus, it would be difficult to distinguish between acute and chronic effects. The Panel suggests for the present time that the Agency conclude acute toxicity tests after a duration of seven days. This will enable the Agency to harmonize international efforts at the OECD level. However, the Panel advises that in the meantime development of a satisfactory standard pupation assay should be a priority. The Panel also suggests that a repeated dose testing method would be most realistic for assessing potential acute effects (*i.e*., effects at the larval stage) and chronic effects (*i.e*., effects at the latter brood stages) of pesticides.

10. Tier I Chronic Toxicity Testing Bees. **Section 4.1.2** of the white paper discusses the status of chronic toxicity tests with individual adult and larval bees. At this time, no formal guidelines have been developed for conducting chronic toxicity tests with either adult or larval bees, although studies with individual bees of various ages are routinely reported in open literature.

a. Please comment on the conclusion that adequate procedures have not been sufficiently developed and validated for assessing chronic toxicity to individual bees in a risk assessment context.

Panel Summary

The Panel believes that social characteristics of honey bees present a challenge to interpreting individual bee toxicity testing to the colony level. Thus, the Panel advises the Agency to consider the OECD (The Organization for Economic Co-operation and Development) 213 study guideline (OECD, 1998). It provides a study design that may help in addressing this challenge.

b. Please comment on the potential use of the 10-d adult worker and 7-day in vitro larval toxicity tests discussed in the white paper for assessing chronic toxicity once these methods are fully vetted.

Panel Summary

The ideal study design for a larval *in vitro* study would encompass the entire active feeding period through pupation, with time to pupation, pupal weight and percent pupation as quantifiable and interpretable measures of pesticide impact. In regard to adult toxicity testing, the 10-day viability for adult bees has the inherent advantage of supporting harmonization efforts.

c. Although 10-day adult and 7-day larval toxicity tests have been proposed, please comment on whether alternative durations of pesticide exposure may be more appropriate for determining chronic toxicity for adult and larval bees at a Tier I screen.

Panel Summary

The Panel recommends that the Agency consider the use of caged brood or broodless bee cups (it is important to note that this is not a standard methodology) to assess the social context of honey bees. In regard to larval toxicity tests, the Panel suggests extending the test duration through the pupation stage.

d. The white paper identifies NOAEC No observable effect effective concentration as the chronic toxicity measurement endpoint. Please comment on the possible use of EC_x values as a measure of chronic toxicity for use in RQ calculations.

Panel Summary

The Panel recommends that whenever possible both NOAEC and EC_x values be used. There are uncertainties and challenges in tying the outcomes of both methods with the protection goals. However, both methods have utility for triggering higher tier tests.

e. Please provide comments on what percent effect would be considered a relevant measure of chronic toxicity for individual bees given the potential compensatory effects which honey bee colonies may exhibit relative to the effects of pesticides.

Panel Summary

The Panel believes that compensatory effects as a bee colony/societal phenomenon are most appropriately addressed at the Tier II assessment level. In regard to a specific percent effect, the EC_{20} (20% difference between controls and experiment to trigger concern) may be a conservative endpoint considering that, one model available demonstrated that (Khoury et al. 2011) a loss of 30% or more foragers daily triggers the death spiral.

f. Although the white paper identifies some measurement endpoints for assessing chronic toxicity to individual bees (e.g., survival), please comment on other endpoints to consider in chronic toxicity studies which the Panel believes are important and the associated study design elements.

Panel Summary

Other chronic toxicity endpoints recommended by the Panel include behavioral impacts such as learning assays, automated recording of motility and behavioral interactions. Endpoints for future consideration include gene expression relevant to detoxification and/or antioxidant status, immunocompetence, and senescence. One research study suggested the use of ribosomal RNA fragments as a biomarker (Johnson et al. 2009). Nevertheless, such studies should be validated and it should be clear how the results of these studies can be used in risk assessment.

g. ***Section 4.1.3 4.1.2.1.2.*** *of the white paper discusses the uncertainties associated with developing risk assessments based on studies of sublethal effects when sufficient linkages have not been developed to understand how the sublethal endpoints may be quantitatively related to typical assessment endpoints (i.e., growth, impaired survival, and reproduction) at the whole colony level. Please comment on the proposal to use data on sublethal endpoints to qualitatively (i.e., no Risk Quotient is derived) characterize effects and risk until sufficient linkages between these endpoints and the defined assessment endpoints have been developed (e.g., Adverse Outcome Pathways).*

Panel Summary

The social nature of some bees contributes to colony function as a "super-organism" and adds complexity to the assessment of sublethal endpoints. Nevertheless, it is still important to identify and assess sublethal effects in order to gain insight into potential impacts to the super-organism. The Panel advises the Agency to advance this effort by placing a high priority on generating research data on bee neurotoxicological symptoms. This may provide knowledge regarding the potential for impacts on bee behavior and other sublethal effects. The outcomes of these studies can trigger the need for additional tests at the Tier II level.

11. Tier II Semi-field Effects Assessments (Whole Colony). For Tier II assessments, **Section 4.2** of the white paper identifies two types of test methods that may be used to assess colony-level effects, *i.e.*, semi-field tunnel tests [OECD 75; EPPO 170]; and semi-field feeding studies (OECD, 2007; EPPO, 2001). These studies are intended to help characterize risks identified in the Tier I level assessment that are based on exposures and toxicity data for individual bees and quantified using Risk Quotients.

a. *Although colonies are typically confined to enclosures for Tier II studies and these enclosures can limit the environmental realism of the study conditions, tunnel studies provide an opportunity to collect colony-level effects and potentially exposure information. Please comment on the relative strengths (e.g., foraging activity by adult worker bees is limited to treated crop; trophyllaxis enabled) and limitations (e.g., limited study duration, smaller colony sizes, reduced forage area) of these methods.*

Panel Summary

The Panel believes that these proposed Tier II tunnel studies are very important for providing information at the colony level that LD_{50} measurements made for Tier I assessments are unable to adequately accomplish. An inherent strength of proposed Tier II studies is that they provide the basal level in which to properly quantify LD_{50} estimates. An example of a limitation of Tier II tunnel studies is that it is difficult to acquire a sufficient sample size due to the inherent variability between colonies. Thus, there are increased risks of Type II errors (false negatives) with these high-tiered tests. In regard to feeding designs for these studies, the Panel recommends that feeding with spiked food should only be used to extend exposure when it is necessary and when it is not possible to do a design with extended flowering of the realistic crop (e.g., if there is concern for chronic long-term exposure).

b. Please comment on any other types of colony-level studies that should be considered as part of Tier II.

Panel Summary

The Panel suggests that the Agency consider observation hives to record behavior and estimate interactions as an alternative study design. This will be useful in addressing sublethal effect questions. The Panel also recommends that the Agency explore potential *in vitro*, whole mini-colony tests that could be implemented in the short term that blur the lines between Tiers I and II. For example, Tier I tests of "bee cups" each containing a queen could be confirmed with Tier II tests of "baby nuc" colonies housed in incubators. While still *in vitro*, such colonies would contain queens, brood, comb, and other full-colony phenotypes that could be measured. For all tests that entail artificial feeding, the Oomen Test can be used as an intermediate test between Tier I and Tier II (Oomen et al., 1992). Some special study designs can be employed to investigate special effects or exposure routes. Examples may include studies to evaluate the effect of a pesticide on queen rearing, (e.g., a shook swarm") or "swarm box" method of queen rearing, exposure via the crop and observation of the number of reared queens and rearing success of hatched queens. Such studies would be important since there are some published reports suggesting that queens may be sensitive to some pesticides. The Panel recommends that an alternative study design be developed to investigate such specialized impacts and potential pesticide exposure routes that may cause them.

c. Please comment on the most important endpoints that should be measured in the Tier II studies (e.g., adult forage bee mortality, brood development, queen fecundity, overall colony strength) that are linked to assessment endpoints and their associated protection goals.

Panel Summary

The Panel suggests the following as the most important measures that should be conducted to link assessment endpoints and their associated protection goals:

- Measurement of effects on different colony phenotypes.
- Assessment of the association between foraging activity and colony phenotype.
- Implementation of tunnel studies whose design is based on the expected exposure routes.
- Assessment of the behavior of foragers both at the hive entrance and during foraging on a crop.
- Employment of the OECD 75 method using maps or digital imaging of the brood to assess the impact on queen fecundity by evaluating the amount of brood in the hives at certain intervals, and the effects on larval development and brood rearing success (OECD, 2007).

*d. **Section 4.2.2** of the white paper discusses a full-field feeding design. This methodology is discussed under Tier II assessments since the colony is relatively confined to foraging on either spiked sucrose solutions or spikes pollen. The intent of this methodology is to ensure that colonies are exposed to known residue levels over longer durations than the semi-field tunnel study designs. A limitation to the study is that bees may simply store spiked food rather than consume it and that the reliance on a single source of food may introduce confounding effects (e.g., nutritional deficits) into the study. Please comment on the environmental realism and utility of full-field feeding studies as a line of evidence in characterizing risk to honey bee colonies.*

Panel Summary

The Panel believes that full-field feeding studies are not appropriate for risk assessment purposes. At best, they can be used as supplemental information of limited value as they are not environmentally realistic. The Panel advises that OECD 75 and EPPO 170 guideline studies are used for Tier II and Tier III studies, since these studies may be the most appropriate test designs for capturing contact and oral exposure pathways (OECD, 2007; EPPO, 2001).

*e. As discussed in **Section 4.3.4** of the white paper, it is important to consider the biological significance of a measured effect in addition to its statistical significance. Please comment on the nature and magnitude of effects that would be sufficient to conclude biologically significant effects on the colony and/or the need to transition to Tier III assessments.*

Panel Summary

The Panel provides the following list of prioritized potential measurement outcomes (Table A) as a means of orienting risk assessors towards more-informative measurements that have greater penetration (*i.e.*, they likely measure a stronger biological signal in the system):

Table A. Prioritized potential measurement outcomes

Relative utility for measuring risk	Ecological measures	Colony measures	Individual measures
Higher utility	Mortality in the crop	Colony strength	Pesticide residues
	Mortality at the hive	Weight of the hive	Longevity
	Mortality of drones and pupae	Food stores	Behavioral abnormalities
	Foraging activity in the crop	Presence of the same queen	
	Returning foragers	Ability to requeen	
Lower utility		Levels of disease Over-wintering success	

12. Tier III Effects Assessments. **Section 4.3** of the white paper discusses the proposed risk assessment process in Tier III that relies on assessing effects at the colony level where colonies are not confined (*i.e*., full field) and exposure is intended to represent environmentally realistic conditions. As discussed in the white paper, interpretation of the biological significance of colony-level effects can be challenging, regardless of their statistical significance. Conversely, high variability in field studies can limit the statistical power of the study to detect treatment effects. The paper identifies uncertainties associated with the extent to which bees forage on the treated crop, the size of treatment site, controlling for alternative forage/pesticide use area, and ensuring suitable controls; these factors combine to render these studies highly resource intensive.

a. Please comment on the strengths and limitations of full field studies described in the white paper.

Panel Summary

Tier III studies are intended to be the best test of real world exposures and possible effects. Nevertheless, these studies have many weaknesses since their reliability depends on performing a controlled experiment in an open environment with adequate controls. The European Union has extensive experience in conducting these experiments. When they are correctly conducted they may provide the best way of assessing possible pesticide risk to *Apis mellifera* colonies and other managed pollinators (*e.g., Bombus* or *Osmia*).

Examples of Tier III study strengths:

- ”Natural” exposure can be achieved by allowing bees to forage freely on a treated crop.
- Measurement of endpoints that test for effects at the colony level can be made that are unattainable at Tier I and II level testing.
- Exposure can accurately simulate the actual use of the pesticide.

Examples of Tier III weaknesses

- Variability among colonies may be high thus there is a need for replication and large sample size.
- Study site selection is critical since the presence of untreated crops and vegetation that serve as alternate food sources for bees can result in low to no exposure to the target crop, especially if the plot size is limited or the crop is marginally attractive.
- Plot size will vary with less attractive crops needing larger plots to insure adequate foraging and exposure to the target crop.

*b. Please comment on the proposed modifications to the field study design elements presented in **Section 4.3.2** of the white paper.*

Panel Summary

A concern about Tier III study design that the Panel notes is the lack of control regarding bee foraging which consequently causes uncertainty about pesticide exposure. The Panel recommends lessoning this concern, by performing an estimated calculation of foraging activity (given the Agency identified foraging range of 78.5 mi^2 area) and also by analyzing a relatively small number of samples of foraging bees over the duration of the study in order help quantify the pesticide exposure. The Panel also recommends using *Bombus sp*. or *Osmia sp*. as test species since they fly shorter distances, and consequently, their level of exposure can be greater and/or more controlled.

c. Please comment on factors that should be considered in evaluating the biological significance of effects measured in full-field studies in relation to the proposed assessment endpoints and related protection goals.

Panel Summary

The collective set of the weight of evidence built in the different tiers including information on acute mortality of adults and larvae as well as colony development and behavior should be taken into account. In particular it is important to note that effects on queen survival or replacement are of biological significance and have implications for colony survival that may not be accounted for if test duration is only a few weeks or months. Ultimately it should be recognized that colonies are complex super-organisms and that several endpoint measures may be more appropriate than relying any single endpoint.

d. Please comment on factors and methods that should be considered when extrapolating observed effects at the colony level in semi-field and field studies to those expected to occur in the environment (e.g., spatial and temporal scale of exposure, hive management practices, presence of multiple chemical and non-chemical stressors, etc).

Panel Summary

Tier II and III studies are unable to incorporate all seasonal and colony management practices that may impact the response of colonies to pesticide exposure. However, an adequate level of confidence in a study can be achieved if the study is performed in the normal growing season in which the test pesticide is expected to be used. Examples of factors to consider in conducting Tier III studies include the length of the study (as the longer the study duration the more likely outside variables such as pests and disease will interfere with the results). Also if a crop is major and widely planted, the study may not reflect the level of exposure bees may receive once the pesticide is used widely across a region. The acreage treated may be proportionally greater than in the study (*e.g.*, corn or soybeans in the Midwest).

*e. A number of study design elements are discussed in **Section 4.3.5** of the white paper; however, even in the best designed studies, there can be confounding effects which can limit the utility of these studies in risk assessment. Please comment on factors that should be considered in determining the utility of field studies for pesticide risk assessment, including a discussion of the representativeness of a study for a National Level assessment (i.e., the pesticide may be used anywhere in the United States and its territories).*

Panel Summary

The single measure that gives the highest level of confidence in the study's validity is the measurement of exposure linked to incoming pollen during the exposure period. Additional considerations include: 1) field testing which involves use of formulated products, 2) representativeness of various crop environments in which a compound may be used, 3) size of the test plots (the larger the test plot the better; one Panel member recommends a field plot size of 40 acres.), and the number of bees in a test trial (the Panel recommends that a good rule of thumb for the maximum number of bees in a field trial is 30,000 adult females per acre).

13. Risk Estimation. Sections 2 (Proposed Risk Assessment Process) and **5** (Risk Characterization) of the white paper indicate that the proposed risk assessment process is intended to be tiered and iterative. As part of the Tier I screen, a number of iterations can be conducted on exposure estimates that allow the risk quotient (RQ) values to be further refined and potentially pass the screen without requiring higher tier effects testing at the semi- or full-field level. However, while the proposed Tier I process for bees is quantitative and results in an RQ value which can in turn be compared to a Level of Concern (LOC), higher tier refinements are used to qualitatively (*i.e.,* no RQ derivation) determine whether whole hives will be adversely affected from the use of a pesticide based on the initial screening-level assessment.

a. Please comment on the use of data on individual bees to transition to higher tier studies given that the Tier I studies focus on survival as the primary measurement endpoint although additional endpoints may be forthcoming as test designs continue to develop.

Panel Summary

Higher tier studies are vital to determine pesticide risks under the conditions that bee colonies experience in the field. Measurements of individual endpoints (e.g. LD_{50}), in Tier 1 may not translate effectively into what will happen at the super-organism/colony level. Mathematical modeling (after calibration and validation) can be helpful towards this translation. It would also be helpful to assess pesticide mode-of-action in order to qualitatively assess exposure risks both individually and cumulatively. Although results of a Tier I study may not characterize the overall impact on the super-organism, these studies are critical. When they are appropriately conservative, Tier I tests trigger higher tier studies that have the potential to detect adult or larval toxicity and behavioral abnormalities via multiple exposure routes.

b. Please comment on the derivation of the Level of Concern (i.e., LOC=0.4) and the extent to which it serves as an appropriate screen to transition to higher tiers of testing/refinement.

Panel Summary

EPA proposes to set this LOC value at 0.4 which is equivalent to approximately 10% mortality. Considering the uncertainty associated with extrapolating effects on individual bees to effects on the colony, it is hard to consider an appropriate level of concern for adult bee mortality at the Tier I level. Other endpoints such as behavior, development, and growth rate may be more appropriate and reflective of colony health. Currently, the science does not support the premise that 10% mortality of adult bees would result in colony collapse or even instability. Thus, the Panel believes that the 0.4 LOC value may be too conservative. To estimate a more realistic mortality, the Panel suggests that the Agency consider exploring the use of a sensitivity analysis of mortality in currently existing population biology models such as described by Fefferman and Starks (2006).

c. Please comment of the quantitative aspect of the screening-level (Tier I) assessment and the use of Tier II and Tier III whole hive studies to qualitatively characterize risk.

Panel Summary

Generally, the Panel agrees that higher tier studies are best suited to assess real world scenarios. The conservative approach of Tier I tests is to ensure that higher tiers are triggered when appropriate. The Tier II tests represent a worst case scenario, and are sensitive to finding effects (since they involve bee confinement scenarios). Tier III studies embody more realistic situations in that they take into account all multiple stressors and multiple issues that cannot be addressed by Tier II studies (e.g., homing behavior).

d. Please comment on the assumption that the effects on individual bees measured in laboratory studies must be considered in the context of whole colony studies conducted under semi-field and full-field conditions.

Panel Summary

The Panel believes that individual phenotype is not an appropriate substitute for colony phenotype. Nevertheless knowing the impact to the individual bee is critical since colonies are made of individuals. Considering the type of effects and the nature of the pesticide, higher tiered studies may need modifications to address the chemical under evaluation.

e. Please comment on the proposed use of a weight-of-evidence approach based on information obtained from multiple tiers of risk assessment for characterizing pesticide risks to honey bees.

Panel Summary

Identifying a line of evidence with strong linkages among its pieces is fundamental to the weight of evidence approach. The pieces of evidence the Panel recommends utilizing are: 1) reliable and repeatable high throughput methods, 2) properly investigated incident reports, and 3) qualitative evidence (which is more important in social systems than simple mortality derived from solitary systems).

f. Please comment on how best to characterize overall uncertainty or weigh different areas of uncertainty in risk characterization.

Panel Summary

Tier II and Tier III studies are prone to Type I (appearing to have significant effects, when in reality there are none) and Type II (appearing to have no significant effects when these effects do exist) statistical errors respectively. To minimize uncertainties, the Panel advises that results of these tests be validated against the preponderance of evidence discovered at the lower tiers.

g. Please comment on how to focus/prioritize uncertainties when designing and interpreting Tier II and Tier III studies.

Panel Summary

The Panel recommends focusing on the following areas when designing and interpreting Tier II and Tier III studies: 1) increasing the sample size, 2) including replication, 3) factoring in genetic variability, and 4) concentrating on the Tier II level. The Panel also advises the Agency to consider utilizing Bayesian methods in interpreting data as these methods are intended to determine "what is the significance of what we were able to observe" rather than applying a "significant or not" test that relies on traditional views of statistical power that may be unrealistic for studies of expensive or difficult scenarios.

14. Colony-level Modeling. As part of the proposed risk assessment process, **Section 5.4** of the white paper discusses the concept of using colony-level models as a means of integrating exposure and effects information generated from the multiple risk assessment tiers and in turn linking this information quantitatively to the proposed assessment endpoints. Conceptually, such models could inform the need for transitioning from lower tiers to higher tiers in the risk assessment process. They could also be considered in identifying critical design elements of higher tier studies (e.g., semi-filed or full field studies).

a. Please comment on the concept of using colony-level ecological models to inform the proposed risk assessment process for honey bees, as indicated above.

Panel Summary

Generally the Panel believes in theory that utilization of colony-level models should enhance the capacity of performing valid risk assessments for honey bees. The Panel cautions that there are many potential drawbacks to models. Most notably it would be a mistake to operate under the presumption that a model that incorporates all aspects of colony life should be used to inform ecological risk assessments. Such models would inherently have a large amount of uncertainty. Instead, models that concentrate on smaller more localized questions should be utilized to support risk assessment.

The Panel recommends that the Agency consider "analytical" models in addition to the simulation models being considered. An analytical model is a model that focuses on the causal relationships among variables and parameters. Equations are used to capture these relationships, and provide insight into system-level processes (e.g., threshold behaviors for phase transitions). The Panel advises that it is important to validate models that are used via a statistical analysis of the correlation between model outcomes and measurements made in real-world scenarios. The Panel cautions against using modeling outputs to fill data gaps.

*b. Please comment on the state of the science regarding available honey bee models discussed in **Section 5.4.2** of the white paper in relation to their potential application in a regulatory risk assessment context. In particular, please comment on the extent that such models have been evaluated using empirical data related to honey bee population dynamics and the availability of such data for their parameterization.*

Panel Summary

The Panel notes the limited availability of input data for many of the models discussed in the white paper. It is important to emphasize that more input (*i.e.* measurements of parameter values and valid characterizations of functional relationships) and validation data (*i.e.*, measurements for comparison with predicted model outcomes) are needed to support models. Consideration must also be given to what is considered an appropriate level of agreement between measured and predicted values during both model calibration and validation. While the Panel agrees that it will be important to collect further input data (i.e. measurements of parameter values and valid characterizations of functional relationships), and to collect validation data (i.e., measurements from the real world for comparison with predicted model outcomes), the Panel notes that analytic models can be useful in the absence of such data. The Panel cautions that even when the proper input data and validation data are available there is a danger in assuming that a model consistently and accurately predicts one outcome of a system without a peer review step for the analytic representation of relationships among model components. The Panel believes that few, if any, of the models described in the white paper have been compared to other models or tested for robustness of model choices apart from their specific focal questions. The Panel also advises that a model under consideration undergo appropriate sensitivity and robustness testing for individual parameters, variables, and functional relationships. One panel member stresses that although models may have the potential to inform risk assessment, the state of modeling is currently insufficient for practical use and that instead there should be more reliance on higher tier tests.

c. Please comment on the most important elements that should be considered in reviewing available honey bee colony ecological models for potential application in risk assessments.

Panel Summary

The Panel concludes that there should be two initial questions with any model. They are: 1) "What insight does it provide that accurate empirical studies do not?" and 2) "How does the model spare effort or expense to provide those insights?" Furthermore, the Panel concludes that the models introduced in the white paper are not designed to answer these types of questions. However, they may be helpful in developing answers within each level of concern and in guiding the risk assessment.

DETAILED PANEL DELIBERATIONS AND RESPONSE TO CHARGE

1. **Section 2.2.1** of the white paper discusses the protection goals and associated assessment endpoints for assessing risks to honey bees (*Apis mellifera*). The protection goals are:
    - protection of pollination services;
    - protection of honey and hive product production; and,
    - protection of pollinator biodiversity.

    As described in the white paper, assessment endpoints are based on their ecological relevance, their susceptibility to known or potential stressors and their relevance to protection goals.

*a. Please comment on whether the assessment endpoints (e.g., population size and stability of managed bees, quantity and quality of hive products, and species richness and abundance) identified in **Table 1** in **Section 2.2.1** of the white paper are consistent with the Agency's protection goals. Please include a discussion of any additional assessment endpoints that may be necessary to meet those protection goals.*

The Panel commends the Agency for doing an excellent job in writing the white paper. It is a very thorough document. It is apparent the white paper was written with a great amount of forethought. The Panel believes that the Agency's proposed Pollinator Risk Assessment Framework" is a step in the right direction for improving the assessment of pesticide risks to pollinators. To further enhance the Agency's effort, the Panel recommends the following regarding the assessment endpoints identified in Table 1 in Section 2.2.1 of the white paper.

In column 2 row 1 of Table 1 of the white paper, the term "managed bees" is used. The Panel notes that the Agency should be clear in identifying managed bees as all commercial bees including commercial bumble bees, alfalfa leafcutting bees, mason bees, and any other bee species that may become managed in the future.

The Panel is unclear regarding how the Agency defines the term "population size" listed in Table 1. The Panel recommends that the Agency define the term population size to specifically mean "the numbers of colonies."

The Panel agrees that it is appropriate to identify "quality and quantity of hive products" as assessment endpoints in Table 1. However, the Panel believes that it may be adequate to use only hive product quantity as an assessment endpoint. This would avoid confusion since hive quality can mean many things that may not be relevant to assessing pesticide risks. However, quality of hive products would be an important assessment endpoint, particularly if it were solely assessed on the amount pesticide residues in a hive.

The Panel believes that the Table 1 (row 3) protection goal, "contribution to pollinator biodiversity" is too broad and unrealistic. The assessment endpoint of this goal, species richness and abundance, is also too broad and unrealistic. Furthermore, there are no measures listed for species richness and abundance throughout the Agency's white paper.

The Panel believes that this broad protection goal is not realistic because pollinators are comprised of large numbers of organisms that include not only bee species but many other

pollinating insects (e.g. some butterflies, pollen wasps, pollen beetles, etc.) and other animals (i.e. some birds, and bats). Bee species native to the Americas are also vastly diverse. There are approximately 4000 native bees in North America, and new native bees are discovered every year. Notably, "honey" bees (including the entire genus of *Apis*) are not native to the Americas. Given the large numbers and diversity of pollinators and given the specific scope of the Agency's white paper, the Panel recommends that the Agency adjusts this protection goal to be "contribution to bee diversity."

b. Please comment on whether the measurement endpoints at the level of the colony (e.g., colony strength and survival, contamination of pollen and nectar and species richness and abundance) identified in ***Table 1*** *are consistent with the assessment endpoints identified in the table and any additional assessment endpoints discussed in Part "a" of this question. Please include a discussion of any additional measurement endpoints that may be necessary to represent those assessment endpoints.*

With regard to the measurement endpoint of "colony strength", the Panel notes that approximately 98 % of bee species are solitary. Thus "colony strength" is not an applicable measure for solitary bees. "Effective population size" may be the more appropriate measurement endpoint for those other bee species. However if honey bees serve as a surrogate for all bee species, the Panel believes that this endpoint should be fine as long the Agency is aware of this distinction.

The Panel suggests that the Agency add "colony development" as another measurement endpoint along with colony strength and survival. Colony development would entail the growth of a colony during a growing season, in essence adding a temporal component to the measurement.

The Panel also recommends that the Agency remove the term "quality" but keep the term "pesticide residues" within the category of measurement endpoints at the population level and higher. In addition, the Agency should add pollen, propolis, and royal jelly to the list of hive products where pesticide residues are measured (all of these are sold as commercial products). The measured endpoint will read as "quantity of hive products and residue levels in honey, pollen, wax, propolis, and royal jelly." This recommendation will reflect the potential exposure of native and honey bees to plant resins that may be contaminated with pesticides.

The Panel agrees that species richness and abundance is an important assessment endpoint; however, measurements of this type are not mentioned in the white paper. Additionally the Panel notes it is impossible, by definition, to assess greater species diversity using only one species, (in this case, the honey bee) as a surrogate. This is particularly true since the honey bee is a domesticated organism that is not a recent native to the Americas. Moreover, there is a considerable body of research which suggests that the presence of honey bees—in certain habitats and ecological conditions—may impose a negative impact on native species of bees through competitive displacement. (e.g. Sugden et. al. 1996, Kremen et al. 2002, Goulson 2003, Klein, et al. 2003, Mayer and Johansen 2003, Tepedino et al. 2007). For these reasons, the Panel believes that the honey bee should not be used to measure species richness and diversity. Consequently, the following Table 1 footnote is inaccurate and the Panel advises the Agency to delete it from the Pollinator Risk Assessment Framework:

Table 1 Footnote:

> "Use of honey bees as a surrogate for other insect pollinators has limitations; however, it is assumed that as with all surrogates, data on individual organisms as well as colony-level data would provide some relevant information on the potential effects of a pesticide on both solitary bees as well as eusocial taxa. In addition, protection of honey bees would contribute to pollinator diversity indirectly by preserving the pollination and propagation of the many plants species pollinated by honey bees, which also serve as food sources for other pollinating insects."

The Panel also recognizes that species richness and diversity are good assessment endpoints to address. Nevertheless, the Agency has not addressed how this will be done in the white paper. Genetics population modeling indicates that genetic variation within a species can be a very effective means of determining the long term effects to population size (Laude and Barowclough, 1987; Frankham, 1995). Thus, the Panel suggests that measuring genetic variation within a species may be a potentially useful measurement for assessing species richness and diversity.

c. *Please comment on whether the measurement endpoints at the level of the individual bee (e.g., individual adult and larval [brood] survival, queen fecundity, brood emergence success, worker longevity) identified in* ***Table 1*** *are consistent with assessment endpoints identified in the table and any additional assessment endpoints discussed in Part "a" of this question. Please include a discussion of any additional measurement endpoints that may be necessary to represent those assessment endpoints.*

Individual worker survival is a good individual-level measure. In addition, the Panel recommends the inclusion of larval survival. Larval assays will likely be very important in understanding pesticide effects on bees. The Panel suggests using the term larvae instead of brood so that it is clearly inclusive of all bees and not just honey bees. The term brood ordinarily only refers to honey bees. The Panel also recommends that the Agency also consider evaluating larval development time. This includes the time period from when an egg is laid to when the cell is capped and the time from capping to successful hatch. The negative effects of some pesticides (i.e., insect growth regulators) may be best detected in this sub-lethal measure.

The Panel is unclear how the term "brood success" is being defined. Thus, the Panel recommends changing this term to "larval survival" and "delayed development instead of delayed success" (Row 2 of Table 1).

Brood size is a good measure, and there are fairly standard ways for measuring this in the hive. However, this term may be mistaken to mean "the size of individual larvae." Thus, the Panel recommends renaming brood size to be brood "nest" size and define it as the area of the sealed brood within a honey bee colony.

Queen fecundity is also a very useful endpoint, but at this time, it is very difficult to quantify. Queen reproductive potential (*sensu* Delaney et al., 2011) is perhaps a better endpoint than queen fecundity as the latter generally refers to the egg laying rate of a queen (or the total eggs in a defined time period). Reproductive potential could include the size and quality of the brood nest,

or survival of eggs through the larval, pupae and adult stages.

Adult worker bee longevity can be difficult to measure because it requires taking the bees out of their colony. Removing workers from the colony to use them in bioassays greatly reduces their longevity, and thus may affect the study results. Another method for measuring adult longevity is to mark or tag a large number of workers in a colony and monitor them over an extended period of time (e.g., Rueppell et al., 2007). This may give a more reliable measure of adult longevity, but it is also very time consuming and labor intensive. However, electronic monitoring devices for individual bees can be used to simplify this task.

2. **Section 2.2** of the white paper discusses a series of conceptual models for assessing risks of honey bees resulting from pesticide applications. These models depict the nature of the stressor (*i.e*., nonsystemic and systemic pesticides that are applied as foliar treatments, soil and/or seed treatments, or trunk injection), its source (*e.g*., direct deposition on bees or pollen), the exposure media (*e.g*., residues on plants, residues in/on pollen/nectar), receptors (*e.g*., foraging bees, developing brood) and attribute changes (*e.g*., reduced [bee] population, reductions in the quantity/quality of honey).

a. *Please comment on whether the conceptual models depicted in **Figures 4** through **8** are consistent with the protection goals and assessment endpoints identified in **Table 1** and discussed in Question 1.*

The Agency's conceptual models are comprehensive and provide substantial detail. However, the Panel proposes the following revisions (See Panel proposed revisions in Figures 4-9 below).

All the conceptual models should be modified to include the potential receptor pathway/exposure route via wax and propolis specifically during the larval stage. Wax and propolis are already considered as a potential exposure/receptor route to adult bees. However, they also need to be considered as potential routes of exposure for the larval bees.

Each of the models has dashed lines which represent routes of exposure that are not considered to be major pathways. The Panel advises that the Agency replace these dashed lines with solid lines because these pathways may potentially be significant.

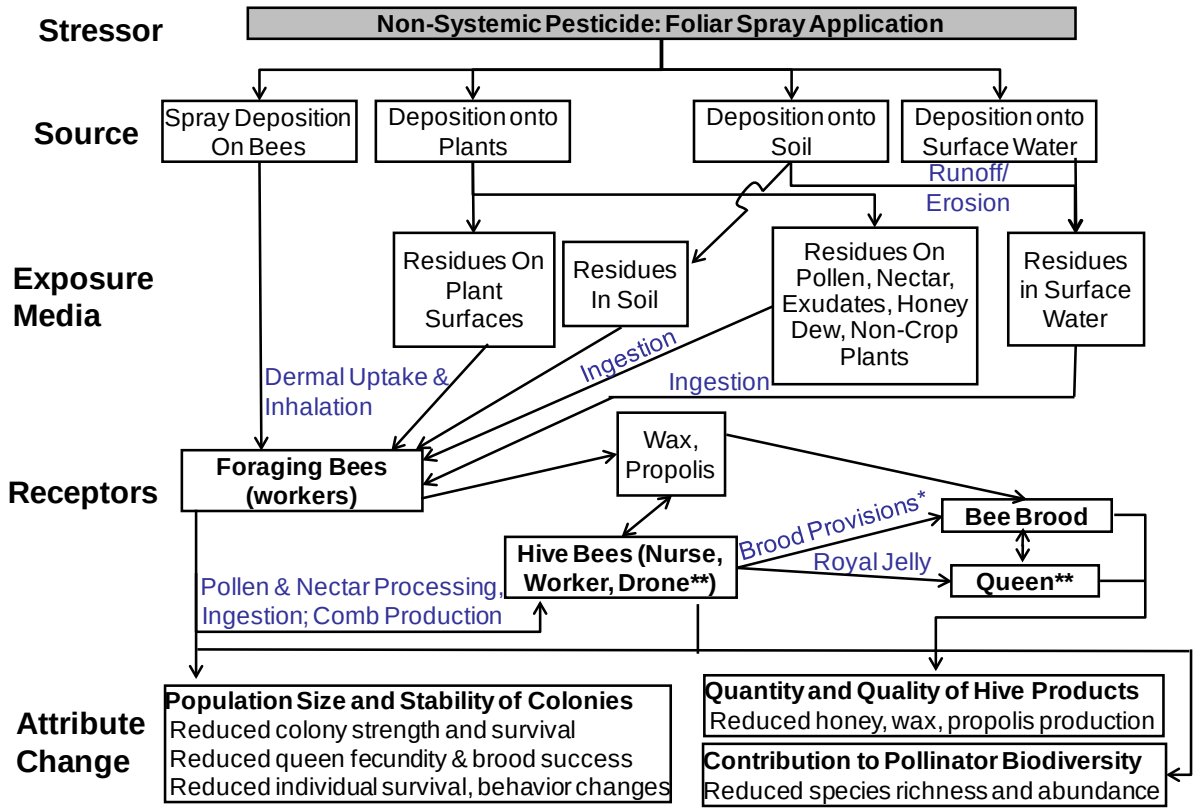
Because honey bees are being used as a surrogate for all managed bee pollinators, exposures that are potentially important to non-honey bees (referred to here as non-*Apis* bees) should be included in all diagrams (even when a specific exposure is not important to honey bees). For instance, many non-*Apis* pollinators are ground nesting and all life stages will have direct contact with soil. As an example, residues in soil may be a significant source of exposure to solitary bees that use mud (made from soil) to partition their nests. When mud is used in the nests, the adult female bee may be exposed to residues while building her nest, and the larvae may be exposed in the nest. Alkali bees, which nest in the soil and are used for alfalfa pollination, provide another example of non-*Apis* exposure via soil. The alfalfa leafcutting bee is also an important commercial pollinator for alfalfa seed. As many, or nearly as many, managed leafcutting bees are moved into alfalfa seed fields (many imported from Canada each year) for pollination as honey bees are moved into almonds for pollination. Adult alfalfa leafcutting bees

can frequently be found crawling on the ground in front of nesting shelters in the morning before the day has warmed sufficiently for them to fly well. Soil exposure routes could also result from wind and tillage generated dusts that settle on plant leaves. These leaves could be used by leafcutting bees as part of the nesting structure, exposing both the nesting females and the larvae that develop inside the leafy nests. Thus, the Panel recommends that soil be included as a significant potential route of exposure in Figures 4, 5, and 6. In addition, topical exposure routes should be added to Figure 7, and a new figure should be created to include non-systemic soil applied pesticides for the same reason.

Abraded seed coatings and dust can drift during sowing (Krupke et al., 2012). Thus, pesticide treated seed should be given consideration as an exposure route. In the spring of 2012 one hundred and thirty bee incident reports associated with corn planting were reported in Canada. In addition many reports were made in the US (although most of the US reports are being analyzed by state labs and have not yet been reported to the EPA). In Indiana all the dead bees from the incidence reports made to the State contained residues of the neonicitinoid insecticide, clothianidin.

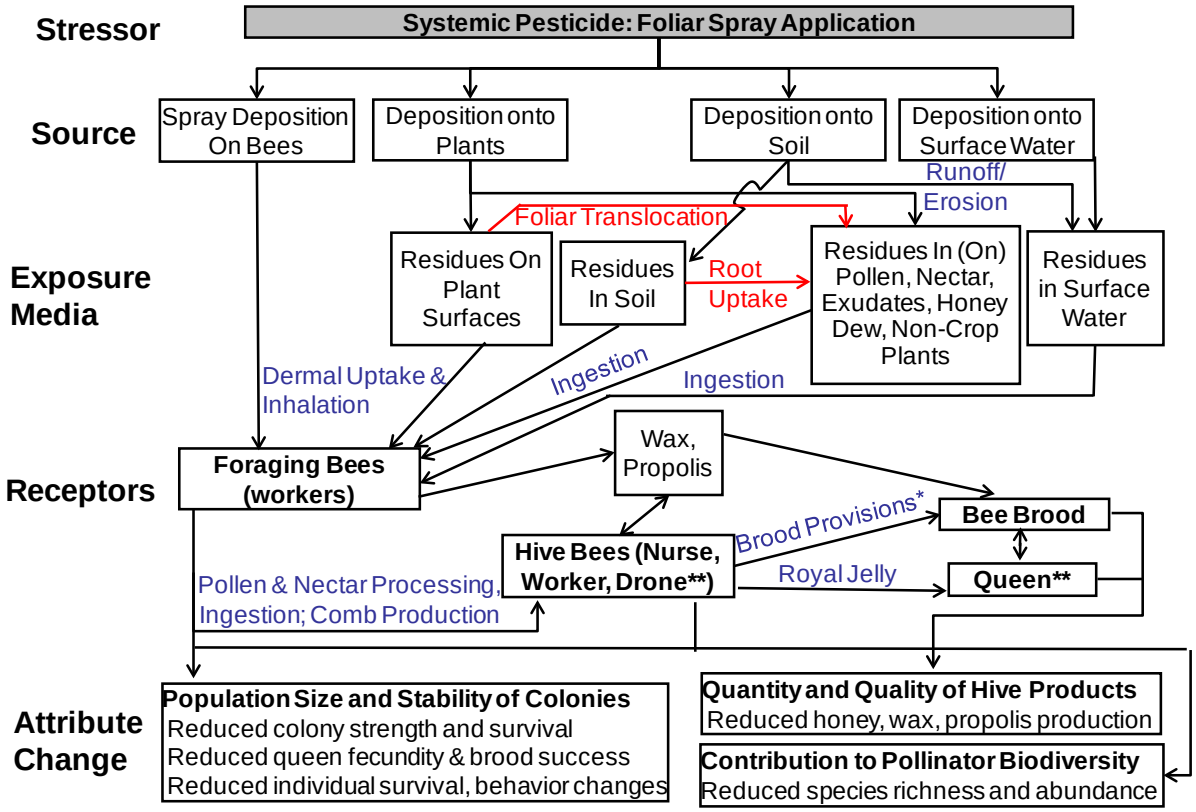
The Panel believes that it would be a good idea to consider summing the larval RQs from oral and contact exposures rather than evaluating the RQ's separately. This idea would mimic the simultaneous acute and oral exposures that occur for the larvae.

Also in regards to bee larvae, the total food consumption during larval stage would be a better assessment of larval dietary exposure. Using the total food consumption would more accurately assess the exposure of honey bee and non-*Apis* bee larvae. For honey bee larvae the period of time over which food is consumed is a short 7 days and food provisions for solitary bee larvae are typically collected in a single day.



* Brood initially rely on royal and brood jelly, but later consume processed pollen and honey; whereas queens rely solely on royal jelly.
** Interception of spray droplets is also a potential route of exposure during mating and orientation flights

Figure 4. Generic Conceptual Model of Non-Systemic, Foliar-Applied Pesticides for Honey Bee Risk Assessment.



* Brood initially rely on royal and brood jelly, but later consume processed pollen and honey; whereas queens rely solely on royal jelly.
** Interception of spray droplets is also a potential route of exposure during mating and orientation flights

Figure 5. Generic Conceptual Model of Systemic, Foliar-Applied Pesticides for Honey Bee Risk Assessment. Red arrows depicts systemic pathways. The red font indicates the Panel's proposed revisions.

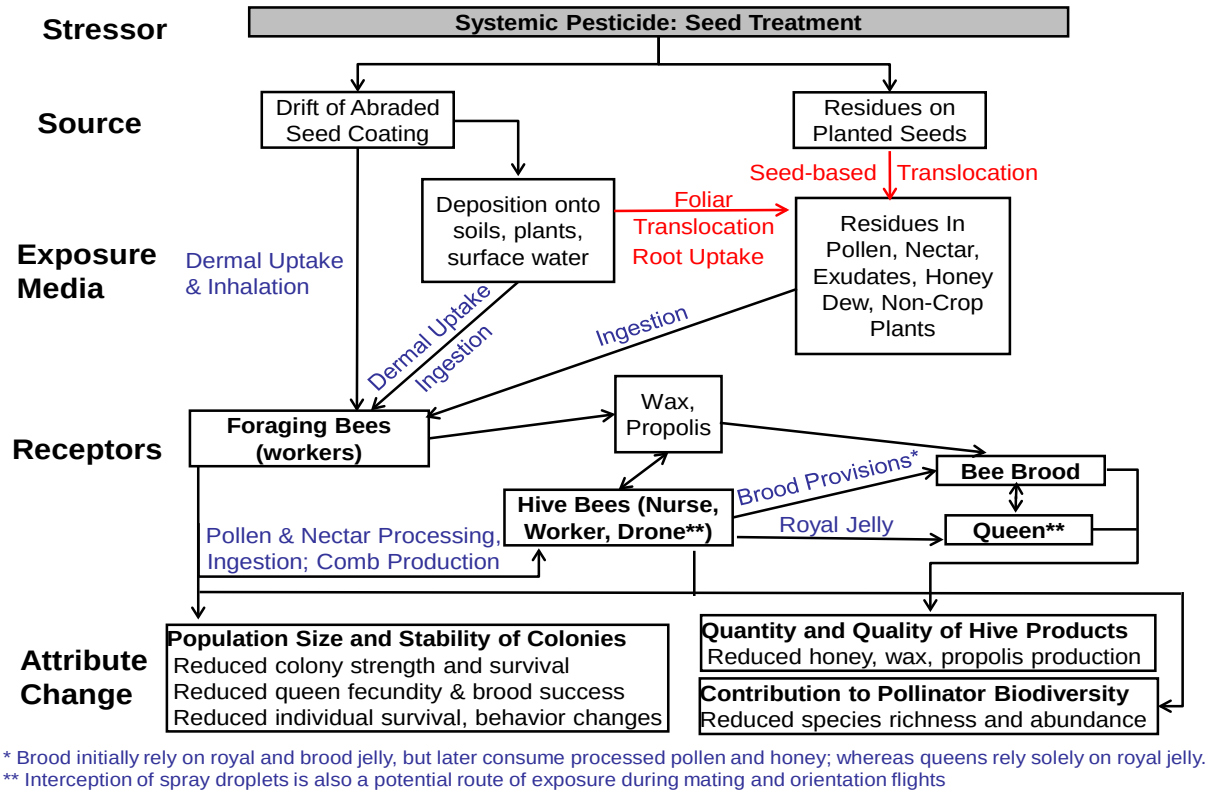


Figure 6. Generic Conceptual Model of Systemic, Seed Treatment Pesticides for Honey Bee risk Assessment. Red arrows depicts systemic pathways. The red font indicates the Panel's proposed revisions.

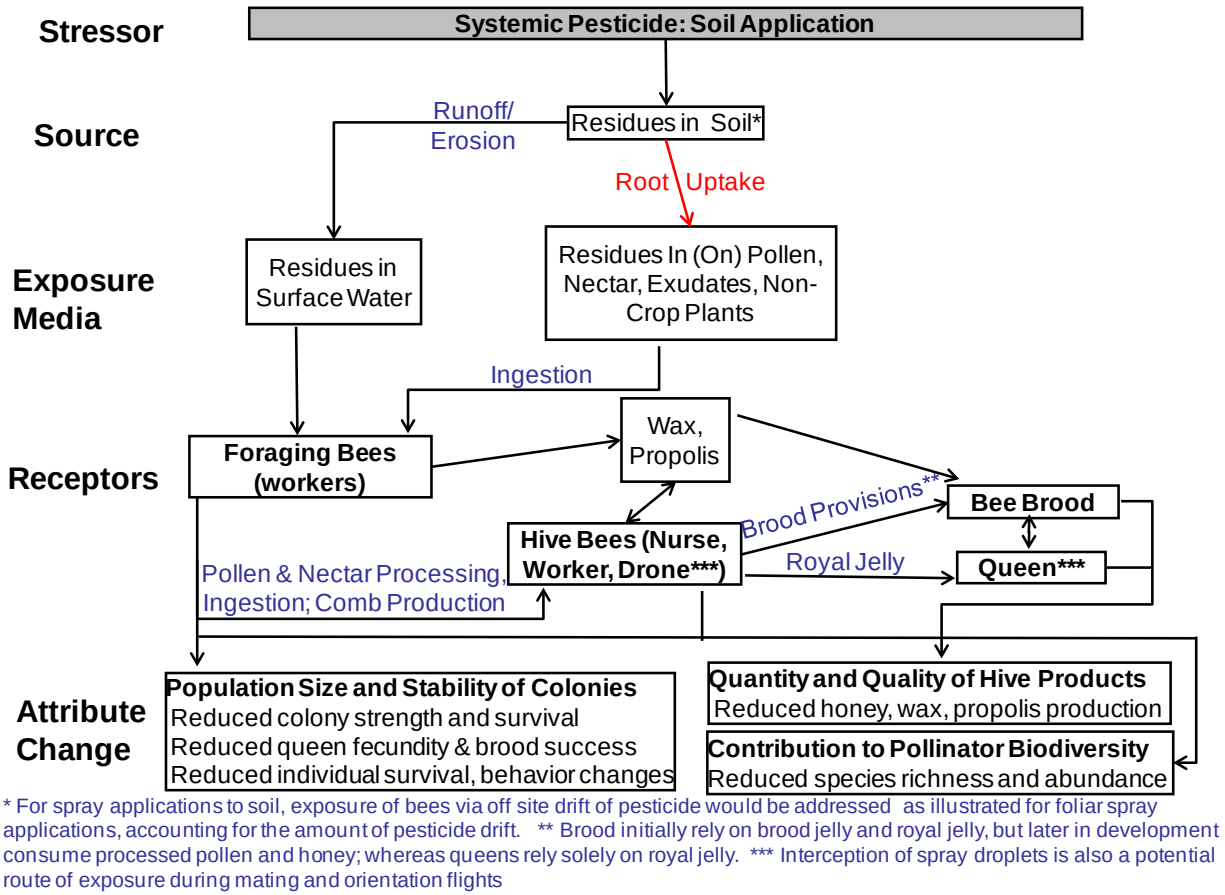


Figure 7. Generic Conceptual Model of Soil-Applied Systemic Pesticides for Honey Bee Risk Assessment. Red depicts systemic pathways. Red arrows depicts systemic pathways. The red font indicates the Panel's proposed revisions.

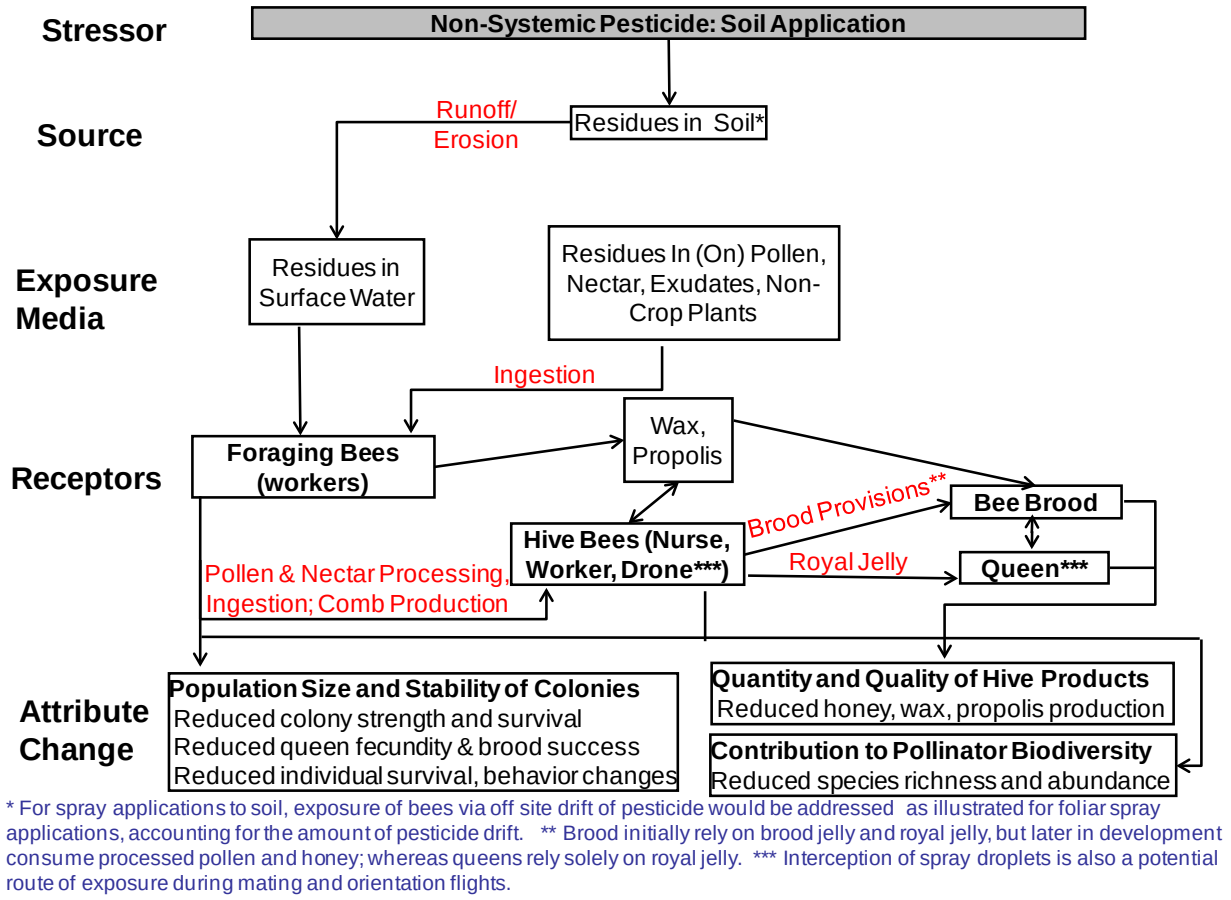


Figure 8: Generic Conceptual Model of Soil-Applied Non-Systemic Pesticides for Honey Bee Risk Assessment. The red font indicates the Panel's proposed revisions.

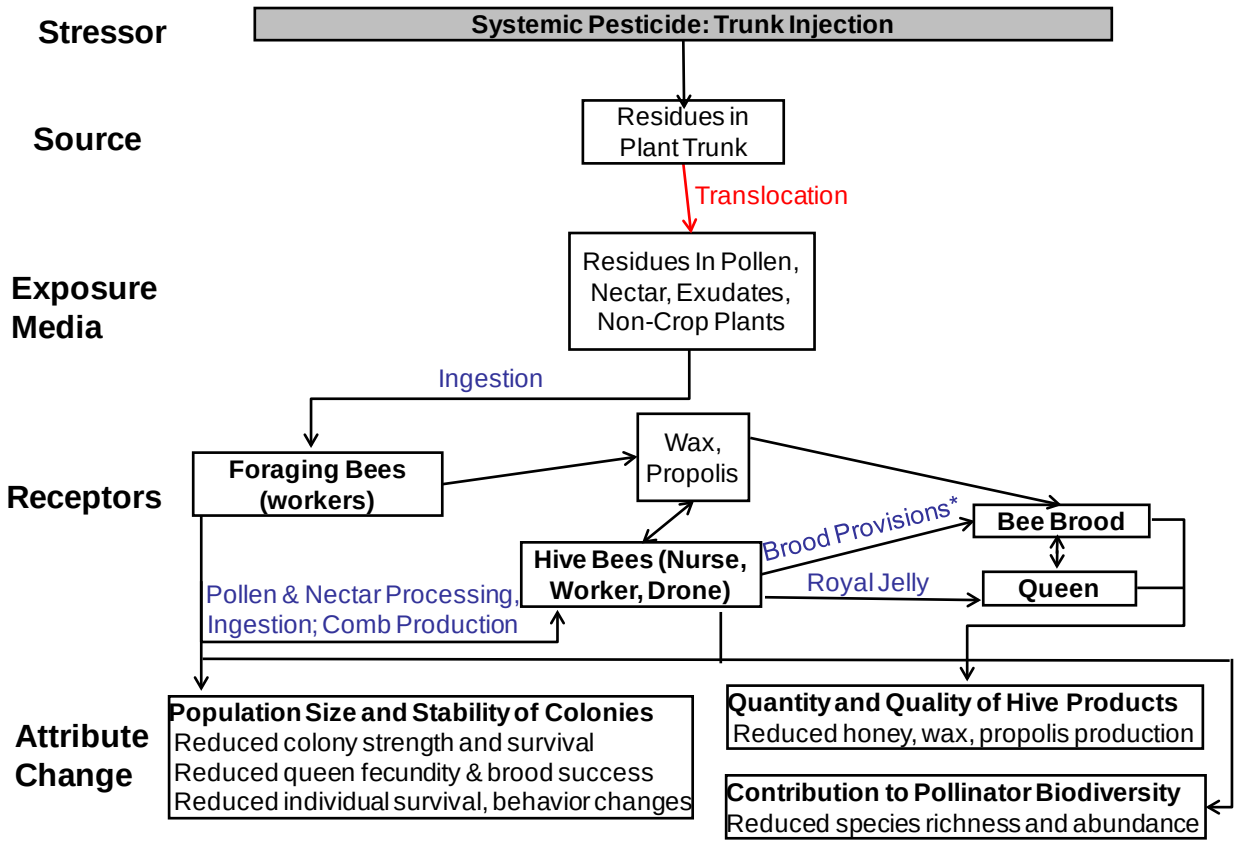


Figure 9. Conceptual Model of Trunk-Injected or Trunk-Drench Systemic Pesticides for Honey Bee Risk Assessment. Red arrows depicts systemic pathways. The red font indicates the Panel's proposed revisions.

A few other potential routes of exposure are missing from the conceptual models. Non-crop plants (weeds, other flowering plants) may be a significant source of pollen, nectar or propolis and may be impacted by overspray. Additionally, not all plants in a field are crop plants; some of these may provide resources to pollinators and the plants may take up residues in soils. This possibility should be incorporated into the models.

The wax and propolis may become contaminated by foraging bees that receive pesticide exposures which may then become a source of repeated pesticide exposure to hive bees (nurses, workers, and drones). This may also lead to one means of exposure to the brood. If the contaminants are persistent in wax, contaminant levels could increase over time. It is also possible, if the contaminant is volatile, that there could be inhalation exposure.

Surface water is assumed in the Agency models to be a minor source of pesticide exposure, but the Panel does not completely agree. There are two types of water used in the hive: 1) water that is consumed by the bees and 2) water that is used for cooling the colony. A significant proportion of the water collected by bees is used for cooling. Any pesticide contaminants in the

water will remain in the hive after the water has evaporated providing a source of future exposure to the bees. Thus, puddle water and water from irrigation ditches should be included in the surface water exposure. Pesticide residues in puddle water and water from irrigation may contribute relatively high exposures to bees. Thus, the Panel recommends that these be added explicitly as potential surface water routes of exposure.

It is true that foraging bees are the main life stage that receives initial exposures, but their only function may be to transport potentially contaminated materials. These bees do not necessarily ingest the materials gathered, e.g., nurse bees consume more pollen than do foragers. Likewise, foragers transport water to cool the hive. Foragers carry propolis and may have 'dermal' contact with residues. But often, it is lifestages within the colony (or nest in the case of solitary bees) that have the greatest exposure rates, or consumption rates, of contaminated materials.

3. The focus of the white paper and proposed risk assessment process on the honey bee reflects two important factors: 1) honey bees are considered one of the most important pollinators globally from both a commercial and ecological perspective; and 2) standardized test methods for evaluating exposure and effects of chemicals in a regulatory context are more developed with the honey bee compared to non-*Apis* bees. Both the Introduction (**Section 1**) and the Problem Formulation sections **(Section 2**) of the white paper indicate that honey bees have historically served as a surrogate for other beneficial insects. As discussed in **Section 5.3** of the white paper, there is uncertainty regarding the extent to which honey bees serve as surrogates for native species, especially where life history strategies and differential sensitivity across species render native species more or less vulnerable to pesticides.

a. Please comment on the extent to which the assessment of risks to the honey bee may serve to meet the protection goals identified in the white paper (i.e., protection of pollination services; protection of honey and hive product production; and protection of pollinator biodiversity).

b. Until guidelines are developed for testing non-Apis species of bees, please comment on the extent to which the honey bee may or may not serve as a reasonable surrogate for non-Apis bees, given the differences in life history strategies and potential different sensitivities to pesticide toxicants. Please include a discussion of which types of non-Apis bees may be particularly well represented by either the individual-level or colony level endpoints identified in Table 1 of the white paper; as well as which types of non-Apis bees may not be as well represented, and therefore, may be the focus of potential areas for future research.

Because subpart Questions 3a. and 3b. are closely related the Panel prefers to respond jointly to them. The Panel agrees that it is appropriate for EPA to use surrogates for ecological risk assessment for all bees and that honey bees serve as a reasonable surrogate. Honey bees vary in their susceptibility to pesticides relative to other bee species;however, this variance is not large. Furthermore, honey bees are easier to obtain than most other bees, and testing methods are more readily available.

If honey bees are used as a surrogate species, it is important that EPA risk assessors and managers recognize that these species are merely representative of other bee species.

In some circumstances differences between honey bees and other bees may need to be taken into account. For example, solitary bees do not have other hive products (i.e. honey, wax, and royal jelly) that are used by people. Also, these bees may be more dependent on pollen for nutrition than other bees, and consequently, they may consume more pollen, or proportionately more pollen than nectar. If pesticide residues are different in pollen than nectar, this could be important for evaluating risks.

Honey bees do not spend time in the soil, or on the soil surface. Therefore, the white paper does not include risk diagrams for non-systemic pesticides applied to the soil, and the pathways do not include contact toxicity for any soil applied pesticides. However, many solitary bees and bumble bees nest in the soil. Others collect soil to make mud partitions in cavity nests. Thus, non-*Apis* bees do come into contact with the soil, and therefore, contact toxicity tests should be included in risk assessments for both the adults and larvae. Honey bees could not be used for these tests as surrogates.

Also, the complex social behaviors of honey bees may buffer them against environmental stresses in ways that other bees cannot. For example, many native bees are specialist pollinators. That is, they collect their food (pollen and nectar) from a limited number of plant species. Non-*Apis* bees also often have much shorter seasons than honey bees. The adults may construct their nests during a short window of time (a few weeks), and after that, the other life stages are contained within the nest until the next season. Thus, all the pollen and nectar collected during this window is what feeds the next generation of bees. Honey bees can store food for long periods, allowing them to avoid collecting nectar and pollen from undesirable sources. Also, honey bees have a highly plastic division of labor among its facultatively sterile workforce, but for solitary bees, all females are reproductive females. Thus, adult mortality could have a greater impact on the native bee populations the next year than for honey bees. For solitary bees, the individual measures are more important than the colony-level measures. Bumble bees form small colonies but overwinter as individual queens (the colonies do not overwinter), so individual-level measures are probably more important for bumble bees, as well.

The Panel suggests that EPA include other bee species for testing as well. The alfalfa leafcutting bee, *Megachile rotundata* and *Osmia* spp. (mason and orchard bees) are the easiest to include for Tier I testing. Both these bees are commercially available in large numbers and would be fairly easy to use for higher tiered tests, as well. Bumblebees (*Bombus* spp.) are also available commercially, and considerable research is available on how to raise them. They would be useful for Tier II tests, although it is more difficult to assess their larval development, brood-nest size, and larval mortality. The Panel recommends that EPA consider using at least one additional bee species other than *Apis mellifera* to address the goal of protecting diversity.

4. **Contact Exposure.** The exposure characterization of the white paper (specifically, **Section 3.1.1.2**) proposes a screening-level (Tier I) approach for quantifying contact exposure to foraging bees for pesticides applied via foliar spray. This proposed method is based on the maximum of residue values on honey bees from a field study conducted by Koch and Weisser (1997). The white paper also discusses a method based on the T-REX upper-bound concentration for arthropods directly sprayed with pesticides while located directly on a treated field. Although the second method is not proposed for the Tier I exposure assessment for honey bees, it could be used to assess contact exposures to other insect pollinators.

a. Please comment on the strengths and limitations of the proposed approach for assessing contact exposures to honey bees in Tier I exposure assessments (i.e., 2.7 µg a.i./bee per 1 lb a.i./A), which is based on the honey bee specific maximum concentration reported by Koch and Weisser (1997).

Direct contact by foliar pesticide sprays can result in some of the highest exposure levels (other than dietary exposure to foliar spray). Consequently, this could lead to the highest level of exposure to foraging bees. Data for bee contact exposure to key pesticides is lacking, thus a proxy methodology for estimating contact exposure is necessary. To this end, the approach described in the white paper has merit; but, for many reasons, this approach also has limitations. In the absence of data, the T-REX model is an alternative that can help identify sensitive parameters and provide general ranges of residue levels. However, relying on a model when actual data is available is not advisable. Thus, T-REX for non-*Apis* bees is more reasonable than for *Apis* bees. Generally, *Apis* bees have been studied far more and have considerably more data available than non-*Apis* bees. Thus, T-REX may be the best approach for non-*Apis* bees and may help fill data gaps for *Apis* bees.

The Panel identifies the following strengths and weaknesses of the proposed method.

Strengths:

- The surrogate tracer used in the Koch and Weisser (1997) study most certainly was not being used anywhere in the region. Thus there are no possible confounding factors associated with bees foraging in unknown areas where the spray was applied.
- The concentration reported was the highest value (normalized), and thus should be protective.

Weaknesses:

- The study was conducted in 1997. A 15-year-old study is not necessarily poor just because of its age, but follow-up and updating are in order.
- The compound used was sodium fluorescein, a fluorescent dye. Its chemistry may be radically different from any pesticide, and one cannot assume it is an appropriate surrogate.
- Bees were foraging during application and subject to direct spray. This may have contributed to anomalously high exposure estimates.
- Normalization to a single application rate assumed a linear relationship between rate of spraying and rate of accumulation by the bees. This is not necessarily valid.

- The Koch and Weisser (1997) approach is reasonable, but the formulation of the spray for the fluorescein is important. It is unknown whether the fluorescein sprayed behaved the same as fully formulated pesticides since spreaders and stickers, surfactants and other adjuvants are routinely combined in pesticide sprays. These adjuvants have a large impact on the distribution and fate of pesticides sprayed on leaf surfaces. Also, it is likely that flower surface areas differ substantially from grass stems. This is important since cuticular waxes and other components on plant surfaces may impact pesticide behavior. Panel members note that there appears to be a pressing need to answer these and related questions experimentally.
- Reporting pesticide loads in terms of per bee rather than per unit mass (e.g., per mg) is a concern. The Panel believes that this will make for a better transference of measurements. Reporting pesticides on a surface area basis may apply to other areas of the paper as well. One panel member suggested that per unit mass for queen would not be a good choice due to the unique aspects of the feeding and longevity of queens.
- The Panel expressed concerned about studying only one race of bee.
- The consumption rate being proposed was on a per day basis. For larvae, the consumption is not the same over the full development period. Thus, total consumption for larvae may be a better approach than consumption on a per day basis.
- The Panel recommends that integrating mass accumulated over time as a good approach for the adult bees but not for the larvae.

b. *Please comment on the potential utility of the T-REX upper-bound residue value (i.e., 12 µg a.i./bee per 1 lb a.i./A), for a broader number of arthropod species to represent contact exposures to honey bees and to other insect pollinators that are directly sprayed with pesticides*.

As noted in Panel's comments on Question 4a, direct studies of bee exposures to foliar applied chemicals are preferred. In the absence of such data, the Panel believes that a model may be a suitable alternative. However, there is substantial uncertainty. The Panel encourages the implementation of a bee exposure study which will generate these data.

In regard to the potential utility of the T-REX upper-bound residue values, the Panel points out the following strengths and weaknesses of this approach.

<u>Strengths:</u>

- The data used in generating T-REX represent actual measurements on arthropods and are expected to be reflective of what might be expected for honey bees and non-*Apis* bees.
- The arthropod data used to generate T-REX are probably no more uncertain than tracer estimates.

Weaknesses:

- T-REX was developed based on 14 studies using carbamate and organophosphate insecticides. Other pesticides, e.g. the neonicotinoids, have a totally different chemistry, thus whether they behave similarly is unknown.
- The original data used in developing T-REX are subject to some question as to the route of exposure.
- Doses and residues are based on normalization that assumes linear upscaling or downscaling, and this may not be the case.

Given strengths and weakness of the T-REX approach, one panel member suggests a potential alternative for estimating the interception of a pesticide during a spray application to a crop. The approach involved consideration of a standard spray rate of 1 lb $acre^{-1}$ which is equivalent to 11 µg cm^{-2}. Using the reported mean surface area of a honey bee, 2.4 cm^{-2} (Roberts & Harrison, 1999) and assuming 100% coverage of a bee during spraying, the maximum potential exposure can be evaluated by multiplying the surface area times the spray rate. In this case the value is estimated to be 26 µg per bee. Comparison of this value to the T-REX upper-bound residue estimate of 12 µg per bee indicates that T-REX approach would result in assuming an interception rate of about 45%. Interception at this rate appears high for a number of reasons, thus the Panel felt that the T-REX estimate was reasonably conservative. This was based on Panel discussions that pesticide sprays to the environment entail a 3 dimensional intercept of pesticide residues (i.e. interception to the various layers of media/biota including plant leaves, stems, and the ground) thus limiting the amount of pesticide that a bee may intercept. A caveat is that the hairs on bees bodies may enable them to retain spray residues at much higher levels than their body surface area (Mean surface areas were: head, 26.8±0.3 mm^2; abdomen, 79.5±2.2 mm^2; thorax (minus legs and wings), 62.8±1.0 mm^2; legs, 68.3±1.1 mm^2). In conclusion, comparison of the direct spray-surface area calculations to T-REX estimates suggests that the T-REX value may provide a reasonable upper bound estimate. In response to this, one panel member adds that, in terms of surface area, perhaps we should consider only active adsorbing surface area. However, another panel member does not like the interception approach for two reasons: the 3 dimensional nature of spraying and the ability of the bees to retain sprays through little hairs on their bodies.

5. **Dietary Exposure (Consumption Rates).** As discussed in the effects assessment section of the white paper (**Section 4.1**), acute oral toxicity data (LD_{50}) are necessary for adult and larval bees in order to characterize the potential risks of a pesticide. Because these toxicity data are expressed on a dose basis (*i.e.*, µg a.i./bee), it is necessary to convert estimated concentrations of pesticides in food (expressed as mg a.i./kg) into doses. Honey bees fulfill their nutritional requirements through consumption of nectar, honey, and bee bread (pollen/honey). In addition to requiring bee bread and nectar or honey, bees also require royal jelly and brood food (jelly) to fulfill their nutritional requirements. In the proposed approach, it is assumed that exposures through consumption of nectar and pollen are conservative representations of potential exposures through consumption of honey and bee bread, respectively. This approach is likely to be conservative because it assumes that

pesticides do not degrade while honey and bee bread are stored in the hive and have undergone some degree of processing (*e.g*., fermentation). For bees that consume honey, it is assumed that the estimated pesticide exposures can be related back to the original concentration in nectar by accounting for the amount of sugar consumed by bees. It is also assumed that pesticide exposures through consumption of pollen and nectar are protective of pesticide through consumption of royal jelly and brood food given that empirical data indicate that pesticide concentrations in royal jelly are a >100 times lower than concentrations in food consumed by nurse bees. In the proposed approach, pesticide doses received by bees are calculated using nectar and pollen consumption rates for larval and adult worker bees. As discussed in detail in **Appendix 1** of the white paper, the proposed values for larvae and adult workers are based on an analysis conducted by EPA, which built upon work published by Rortais *et al.* (2005), Crailshaim *et al*. (1992 and 1993) and others. For larvae, the proposed total food consumption rate is 120 mg/day, which is based on the total daily consumption of pollen and nectar (based on honey consumption) by larvae during day 5 of the uncapped larval life stage. For adult worker bees, the proposed food consumption rate is 292 mg/day, based on food consumption rates of nectar foraging bees, which are expected to receive the greatest dietary exposures among different types of worker bees. In addition, as discussed in **Appendix 1**, it is likely that these food consumption rates are protective of drones and queens.

a. *Although bee larvae typically consume processed foods in the form of royal jelly and brood food throughout much of their development, they also consume honey and pollen during the last two days of the uncapped period. Please comment on the proposed use of nectar and pollen consumption rates of larvae during the last day of the larval developmental stage. Please include a discussion of the conservatism, strengths and limitations of this approach as well as a discussion of how this value may or may not correspond to data generated from larval toxicity endpoints.*

The Panel generally agrees that unprocessed pollen and nectar will represent a worst case scenario and conservative estimate over processed foods such as brood food or royal jelly. The data presented by EPA on the nearly 100-fold decrease in residues from pollen or nectar into royal jelly supports the conclusion that food processing by nurse bees results in low residue levels in royal jelly, brood food, or both.

Concerning larval assays these studies are conducted using dilute royal jelly but the spiking of this food with specific levels of pesticides does allow for an oral larval exposure. The absence of an artificial diet for larval bees necessitates the use of royal jelly for larval rearing but each lot of royal jelly has to be tested in order to be certain it is free of antibiotic and/or pesticide residues. Recent studies by Crailshiem and colleagues in Austria have demonstrated that larval rearing can be completed and a high level of pupation and adult emergence achieved along with test in observation hives which demonstrated that these *in vitro* reared individuals had equal life-span and foraging rates as colony (*in vivo*) reared sister worker bees (Prevention Colony Losses Network (COLOSS), 2010). Thus, with practice the larval rearing can be used or extended to examine milestones in development such as pupation and adult bee emergence.

The proposed method seems sufficiently conservative except that the Panel questions the Tier I laboratory larval feeding concentrations. The white paper identifies the concentration in royal jelly as being a factor of 100 lower than the original concentration bees ingest before processing, but the Panel recommends using the original estimated adult ingestion concentration because the lower value was derived on too small a number of substances.

The food intake seems to be highly conservative as the value of 120 mg per day is assumed; definitely larvae of the first up to the last larval stages would consume less. This value is even higher than the likely ingestion in the last larval day, when the feeding rate is at its peak. This is confirmed by comparing the value of 120 mg to the amounts used in the feeding regime in the in-*vitro* test of Aupinel which is 50 µL of total food on Day 6 (and 160 µl over the whole larval development), even if the density of the food is higher than 1. It has been demonstrated that with the feeding regime used in the Aupinel *et al*. (2005) method a normal bee size will be obtained.

The Panel believes that it would be a good idea to consider summing the larval RQs from oral and contact exposures rather than evaluating the RQ's separately. This idea would mimic the simultaneous acute and oral exposure that occurs for the larvae. (See discussion of contact exposure in brood cells in Question 2.)

Also in regards to bee larvae, the total food consumption during larval stage would be a better assessment of larval dietary exposure. Using the total food consumption would more accurately assess the exposure of honey bee larva and the exposure of non-*Apis* bees would be more accurately approximated. For honey bee larvae the period of time over which food is consumed is a short 7 days, and food provisions for solitary bee larvae are collected in a single day.

b. *Please comment on the strengths and limitations of basing the Tier I screen for adult honey bees on food consumption rates of nectar foraging bees, including a discussion of the conservatism of this approach, and how it relates to other types of worker bees and castes.*

Tier 1 screening using nectar consumption only is generally appropriate but may not be as valid if newly-emerged bees are used in the Tier 1 assay as these young bees would normally consume large amounts of pollen rather than nectar. Many Tier 1 studies use random aged bees in which case nectar consumption alone for these tests is appropriate.

Queens and drones do not normally feed themselves but are rather fed by worker bees and thus the testing in Tier 1 is acceptable to determine LD_{50} values. However, some consideration must be made as to the high rate of feeding that the queen receives during a day and thus she has the potential to receive a much higher exposure than workers alone. This is in part moderated by the demonstrated lower residue levels found in brood food that is fed to queens. One member of the Panel suggests using protein needs for egg production to estimate consumption rates for queens in a way analogous to the way energetic flight needs were used to estimate consumption rates for foragers.

One simple means to adjust for caste and species differences is to look at exposure as a function of bee weight as opposed to on a per bee basis as is currently done. The chosen value will cover all other bees and seems highly protective for nectar consumption of other casts; as it is

concluded most energy is needed for foraging activity and other casts are also covered. In theory, pollen foragers should also have the same energy demand as nectar foragers. Even if it is not clear if the pollen forager nectar consumption is underestimated or the nectar value is overestimated for nectar consumption, the proposed values seem sufficiently conservative for Tier I. The Panel also agreed that use of consumption rates during 'normal' colony processes may be insufficiently conservative for estimating risk from consumption of contaminated honey or pollen stores during rare but unavoidable circumstances (e.g. nest disruption, smoking, swarming, etc.). Under cases in which bees are expected to gorge on stored food, there may be short bursts of increased exposure from consumption without any difference in the exposure per unit of food and these would be very important to take into account for long-term colony viability.

c. Please comment on the assumption that exposures through consumption of nectar and pollen are conservative representations of potential exposures through consumption of honey, bee bread, brood food and royal jelly, all of which represent processed foods

The Panel concurs that nectar and pollen are conservative representations of exposure levels compared to processed brood food (honey and bee bread) and royal jelly. However, one panel member notes that this assumption is not readily extrapolated to native bees because as larvae they are receiving provisions of pollen and potentially a little nectar. However, native bee larvae aren't fed brood food or royal jelly.

Although the Panel agreed that the processed food would have lower pesticide concentrations, there is some uncertainty regarding the degree to which bee food processing affects pesticide concentrations.

6. **Dietary Exposure.** The dietary exposure methods described in **Section 3.1** of the white paper differ in the nature of the estimated concentrations in pollen and nectar consumed by bees. For foliar spray applications, the proposed approach involves the use of the tall-grass residue value from the T-REX model (v. 1.5) as a surrogate for pesticide concentrations in nectar and pollen. For soil treatments, the white paper proposes the use of the Briggs' model, which is designed to estimate pesticide concentrations in plant shoots resulting from plant uptake of pesticides from treated soil. Estimated pesticide concentrations in plant shoots are proposed as a surrogate for concentrations in pollen and nectar. For seed treatments, the white paper proposes the use of the International Commission for Plant-Bee Relationships' (ICP-BR) 1 mg a.i./kg concentration as an upper-bound concentration in nectar and pollen. The paper explores the strengths and limitations of each method relative to the ability to derive reasonably conservative estimates of pesticide exposures to bees, with a focus on how well the estimates relate to empirically based measures of pesticides in pollen and nectar from crops treated with pesticides.

a. *Foliar spray: Please comment on the analysis presented in Section 3.1.1.1, with a focus on the extent to which the T-REX tall-grass upper-bound residue may serve as an adequate surrogate to represent upper-bound pesticide concentrations in pollen and nectar of flowers that are directly sprayed with pesticides.*

Due to the limited data on actual residue data on pollen and nectar, the Panel believes that the use of the T-REX model was necessary to estimate upper-bound pesticide concentrations in nectar and pollen. The Panel notes the following strengths and weaknesses of using T-REX to perform these estimations.

T-REX Strengths for nectar residues:

- T-REX is a model currently endorsed and used by US EPA in Tier 1 studies.
- The model was developed using experimental (empirical) observations on stems of plants.
- The model uses very few input parameters.
- The nomograph employed uses a large quantity of field data. The upper-limit values are the upper bound of residue values from actual measurements.
- Empirical data for nectar concentrations were compiled to evaluate T-REX upper boundaries. Seven studies were used.
- The T-REX values were compared to several field studies, and the upper bound was selected to be conservative without being excessively restrictive. Tall-grass was chosen because it is 3 times higher than the highest upper bound in the nectar empirical data.

T-REX Strengths for pollen residues:

- Eleven studies for direct spray were used in developing T-REX.
- T-REX values compared to empirical values showed T-REX for tall-grass is 2-times higher than the highest measured pollen value.
- Twenty three data points were employed in the empirical data set.
- The 110 mg/kg (per 1 lb a.i./acre) limit should be protective.

T-REX Weaknesses for nectar and pollen residues:

- Actual nectar and pollen data are limited making this value subject to high uncertainty.
- Normalizing to 1 lb a.i./acre assumes that changes are linear.
- The model does not contain actual values for pollen and nectar from flowers that were directly sprayed.
- The highest empirical values correspond to a mean measurement of pesticides in nectar, not a maximum. Therefore, the comparisons were made to T-REX means not upper bounds.
- Crop specific factors cannot be accounted for.
- T-REX for tall-grass appears conservative, but some published data summarized in the white paper to support this conclusion appear uncertain. In particular a review of the manuscript by Choudary and Sharma identified questions about analytical limits of detection and study design. As indicated this was a field study in which pesticide was applied at 50% bloom with a backpack sprayer. Pollen was collected from bee legs, and nectar removed from stomachs. It is unknown whether bees foraged exclusively in treated areas or had access to flowers that opened after spray application. Further it is

unknown whether flowers and stems received the same dose from the broadcast spray. Table 6 suggests this, but data are limited. Finally it was noted that flower morphology may be important (e.g., dandelion versus tomato blossom) and that solubility may govern nectar concentrations and define upper concentrations. Panelists concluded that relatively simple studies to assess pesticide solubility in "surrogate" nectar (30% sucrose) may provide useful insights.

- T-REX may indeed be conservative when considering only the active ingredient, but it may not be representative of how the formulated product may behave in the environment.

The Panel suggests exploring the managed pollinator coordinated agricultural project that has managed aviaries in several states throughout the US (Mullin *et al*., 2010). This project has collected a significant amount of pesticide residue data on various media including bees, pollen, wax etc. These studies may help the Agency to identify maximum pesticide concentrations on pollen and nectar.

b. Soil applications: Please comment on the analysis presented in Section 3.1.2, with a focus on the extent to which the Briggs' model may generate estimates of pesticide exposure in plant stems that can represent upper bound pesticide concentrations in pollen and nectar of flowers.

There are several strengths and weaknesses associated with using the Briggs' model analysis to predict pesticide exposure estimates in plant stems. They are as follows:

<u>Briggs' model strengths:</u>

a. The Briggs' equation was selected by United Kingdom's Environmental Agency as coming closest to estimating empirical data. Requirements: must be more conservative than empirical but within an order of magnitude.
b. This approximation is also used in other models such as PRZM.
c. All pesticides used in developing the model were non-ionic.
d. The range in log K_{ow} values was -0.57 to 4.6, and this range will include the neonicotinoids.
e. The original model was based only on the log K_{ow} and regression analysis of experimental data.
f. The Ryan modification added three soil parameters to make the model more specific: soil bulk density, soil water content, fraction of organic content in the soil.
g. The resulting model was tested against experimental data on five chemicals. Modeling was conservative to maximize concentrations in solution.
h. Predicted data was within an order of magnitude of experimental.
i. To make the model more conservative, the model was redeveloped by including an extra data point eliminated by Briggs and adding 95^{th} percentile upper bound estimates of TSCF.
j. One of the panel members commented that stem concentrations are conservative relative to pollen.

Briggs' model Weaknesses:

a. It is a model that examines nonionic compounds only. Pesticides that ionize will not be well represented by this approach.
b. Pesticides tested were carbamates (insecticides) and phenyl-ureas (herbicides) and on one species of plant (barley).
c. The soil parameters used in the Ryan modification may not be known, particularly the soil moisture content. Bulk density and f_{oc} are spatially dependent.
d. The model is used for stem transpiration stream concentrations, not pollen or nectar.
e. In testing the model against 5 chemicals with experimental data, f_{oc} was assumed to be 0.01, and bulk density and soil water content were set to a relatively high value. Multiple assumptions were made, and this defeats the purpose of putting in soil specific parameters.
f. Concentrations in stem transpiration stream were assumed to be equivalent to pollen/nectar concentrations.
g. Not all values predicted by Briggs' model were greater than experimental data.
h. Re-modeled data with 95^{th} percentile upper bound did not bring all the data above the empirical data. Therefore, this approach is not fully conservative.
i. An upper bound on water concentration of the pesticide determined by the solubility of the compound in question needs to be considered. It is quite possible to have calculated concentrations that exceed the water solubility, and this should be avoided.
j. Briggs et al is a good first effort. As noted the need exists for data for other plants, and transport of weak acids and bases are not covered by the equation but are impacted by pH both in and out of plant. One possible alternative is the Dynamicrop model which appears to be much stronger mechanistically (Frank et al., 2012). The model is freely available.

The Brigg's relationship was evaluated by Burken and Schnoor with another plant species and a much broader array of chemicals (Burken, & Schnoor, 1998). These data should be part of the Agency's evaluation of the Brigg's model. This will require recalculating the 95^{th} percentile upper bound for the equation.

One of the panel members commented that stem concentrations are conservative relative to pollen. Experimental data show this.

c. Soil applications: Please discuss the relative strengths and limitations of the 1 mg a.i./kg value and the soil uptake model (the Briggs' model) proposed in the white paper as Tier 1 screens, including consideration of the extent to which this method may generate conservative Tier I estimates of dietary exposures to bees. Does the Panel conclude that the one approach may be better suited to specific types of assessment scenarios? If so, please elaborate. Alternatively, if both approaches are equally suited for a Tier I screen, please provide guidance on how to be capture variability and uncertainty in the exposure estimates using the two approaches.

The Panel provided a review of the Briggs' model in the above Question 6b. Thus, it will not be addressed further here (in the response to Question 6c). In the discussion immediately below (Question 6d), the merits of the 1 mg a.i./kg value are discussed and will not be repeated here. Rather, this discussion will focus on the relative merits of Briggs' versus the EPPO approach in the complete absence of empirical data.

The EPPO approach is the default position assuming that nothing is known about the soil, plants, or pesticide. This approach assumes that any attempt to estimate pesticide residues offers no more reliability or unreliability than simply assigning a value.

The Briggs' model approach makes an attempt to integrate information about soil, plant, and chemicals and make a projection of the pesticide concentrations in pollen and nectar. As discussed earlier, the Briggs' approach is riddled with uncertainties and weaknesses, but it does offer a more rational and defendable approach than the EPPO approach. Because the default concentration does not take into account pesticide properties, the Briggs *et al*. approach is much stronger mechanistically. In conclusion, considering the limitations of the Briggs' model and the EPPO approach, the best option is to attain more pollen and nectar pesticide concentrations.

d. Seed Treatments: Please comment on the analysis presented in Section 3.1.3, including a discussion of the strengths and limitations of the use of 1 mg a.i./kg value as an upper-bound concentration for pollen and nectar of seed-treated crops.

Section 3.2.2 of the Agency's white paper states, "if pesticide-specific residues in pollen and nectar are not available, a screening value of 1 mg a.i./kg plant matrix is the assumed exposure (EPPO, 2010)." This assumption is based on the upper limit value from empirical data of pesticide concentrations in different plant parts. The Panel believes the respective strengths and weakness of utilizing this value are as follows:

Strengths

- The white paper states that 1 mg a.i./kg tends to be higher than reported values.
- Most pesticides tend to be 2 orders of magnitude below the selected value.

Weaknesses:

a. Selecting a value is inferior to acquiring experimental data.
b. The value is not completely protective.
c. This value takes into account no aspect of field situations.
d. The Briggs' model is stronger mechanistically.
e. Need more data.

Given the above list of strengths and weaknesses the Panel concludes that if the only goal is to generate a conservative value for Tier 1, this is an acceptable approach. Otherwise, this is a very poor substitute for true data.

e. Please comment on other approaches or data that should be considered for estimating upper-bound estimates of pesticide residues in pollen and nectar as a Tier I screening-level assessment for pesticides applied via foliar spray, soil application or seed treatment.

As listed in the Agency's white paper, the other methods for estimating upper-bound estimates of pesticide residues in pollen and nectar for a Tier I screening-level assessment for pesticides applied via foliar spray, soil application or seed treatment are:

- Direct measurement of the pesticide in question on the plant parts of interest (nectar, pollen, etc.).
- Direct measurement of the pesticide in question in other plant tissues (stems, leaves, etc.).
- Direct measurement of surrogate pesticides in plant tissues.
- Modeling of pesticides in the plant tissues after foliar spraying (T-REX).
- Modeling of pesticide uptake from soil after seed application (Briggs/Ryan).
- Assigning a default value when no other method is available.

Variations of these approaches are available including different models, different ways of designing experiments, different uptake algorithms, and different default values. However, these are simply tweaks to the approach and not truly different.

The Panel believes the Agency needs a data call-in followed by a rigorous evaluation of the data. This was noted in the Pellston executive summary. Much of what has been proposed appears conservative, but uncertainties are high.

7. Consideration of other Exposure Pathways. The proposed measures of exposure are based on what are believed to be the primary routes, *i.e.*, direct contact and ingestion of contaminated pollen and nectar. Additional routes of exposure are considered (*e.g*., dust, drinking water), but not included in the proposed Tier I exposure assessment method. As discussed in **Section 3.1.4.1** of the white paper, effective quantitative screening methods for estimating exposures through contaminated dust are not discernible at this time. The most effective management of bees exposure to pesticides through dust appears to be through pesticide application (*e.g.*, stickers) and seed planting practices, especially since dust exposure is expected to be a concern for only a limited number of pesticides and application scenarios. In regards to pesticide exposures through drinking water, **Appendix 2** presents an analysis of potential exposures to bees through various sources of water to support the exclusion of drinking water exposure in the Tier I screen. The results of this analysis indicate that if bees consume the majority of their water from puddles or ponds, the exposures relative to dietary and direct spray are insignificant. The preliminary analysis indicates that if bees drink a substantial amount of water from guttation fluid or dew, conservative exposures may be similar to or even exceed pesticide exposures through the diet or direct spray. Further investigation concluded that pesticide exposures through dew and guttation fluid are not expected to be as significant when compared to diet, primarily because they are not likely to consistently drink a substantial amount of water from these sources.

a. Please comment on the strengths and limitations of basing the Tier I exposure method on contact and diet. Does the Panel agree that for the majority of pesticide applications, the primary exposure routes for bees will be represented by contact and diet?

The Panel generally agrees that diet and direct contact are likely dominant pesticide exposure pathways. However, the Panel believes that contact with dust generated during planting of pesticide treated seed and granular formulation applications, may be quantitatively relevant. These dusts tend to be strongly enriched with pesticides. Bee mortalities during planting of seeds treated with certain insecticides are widely reported. In addition, the Panel believes that pesticide exposure via contact and consumption of pesticide contaminated water should be evaluated. As reported in the white paper, bees consume water to meet physiological needs and transport water to hives for use in colony thermoregulation. The Panel suggests that risk assessments would be improved by considering dust and water exposures in Tier 1. One panelist notes that substances of concern would be identified in the currently proposed Tier I approach that focuses on diet and contact exposure. Specifically this panelist believes that consideration of dust exposure and effects as well as potential mitigation measures should be considered by higher tier testing and or as a risk management measure.

b. *Dust: If the Panel believes that this exposure route should be quantitatively included in the Tier I exposure method, for the relevant application type(s) (i.e., seed treatment), please discuss the data that may be needed to address the exposure route quantitatively.*

The Panel notes that dusts strongly enriched with pesticide active ingredients can be generated when treated seeds are sown, during application of granular formulations, and by wind erosion of fine particulate matter, in particular PM10 size range material. The Panel agrees with the white paper findings that bees may be exposed via contact to pesticides in dust while in flight or when visiting flowers and other vegetated surfaces where dusts have deposited. Another possible exposure route discussed is contact of dust deposited directly on hives that are downwind of agricultural fields. The Panel acknowledges widespread reports of honey bee mortality associated with dust emissions during planting of pesticide treated seed, and they note that in many cases deaths due to pesticide exposure were confirmed by detection of pesticide concentrations above lethal limits in bees. No reports of bee mortalities due to contact with other dusts including those generated from granular formulation application and in PM10 derived from agricultural fields were noted. One panelist emphasizes that little is known about pesticide levels in PM10 and that more research is needed to determine the types and amounts of pesticides that may be present in PM10 collected in agricultural areas. The Panel notes that a recent doctoral dissertation indicated that PM10 may be enriched in pesticides by factors of 100 to 1000 when compared to nearby surface soil (De Rossi, 2010). A second study indicated that that from 1 to 6 % of the herbicides applied to farm fields may be lost in wide blown dust (Larney et al., 1999).

Regardless of dust source, the Panel recognizes that exposure assessment is challenging noting that there are many variables and factors in quantifying emissions from planters sowing treated seed. They are particularly difficult to quantify because many variables and factors are involved including seed type, seed treatment rate, quality and abrasion potential of coatings, the machinery used and use and efficacy of deflectors and other mitigation measures. To this end the Panel advises that there are several publications that provide insight into pesticide emission rates associated with treated seed planting. For example, the study by Tapparo et al. (2012) indicated that from 0.5 to 2 % of the pesticide applied may be emitted in planter exhaust from

corn seed coated with various commercial insecticide treatments. It is suggested that these data could be used to estimate exposure as follows using an approach similar that described in the Question 4b. response. If a 2% loss rate from the maximum application rate reported by Tapparo et al (1.25 mg of active ingredient seed and 66000 seed ha^{-1}) or 1 kg ha^{-1} (as done by EPA) is assumed then computed emission rates would be equivalent to 17 to 200 ng/cm^{-2}. Multiplying these rates by the reported surface area of an adult honey bee, 2.4 cm^{-2}, (Roberts and Harrison, 1999) assuming complete coverage would provide exposure estimates of 40 to 480 ng/bee. These values are in the same range as estimated in the white paper for dietary exposure from a systemic seed treatment (up to 300 ng/bee per day). It was noted that Tapparo et al (2012) also reported detection of 25 to 1400 ng of the insecticide clothianidin in bees exposed to planter exhaust when corn seed coated with a commercial formulation of this product was sown. Dust exposure may also lead to dietary exposure. Krupke et al. (2012) measured clothianidin levels of up to 88 ppb in bee-collected pollen and most of their measurements were an order of magnitude higher than levels found in pollen from systemic movement of pesticide. Therefore these levels are probably resulting from drifting dust and such empirical measurements provide another means of measuring exposure. Given that dietary and dust exposure estimates are within the same range it appears that consideration of both exposure pathways in exposure estimates may be appropriate (Forster et al., 2012).

c. *Drinking Water: Please comment on the analyses, discussions provided in Appendix 2 of the white paper and the conclusion that pesticide exposure to bees through drinking contaminated water is not expected to be a major route of exposure when compared to contact (following foliar spray applications) and diet. If the Panel believes that this exposure route should be quantitatively included in the Tier I exposure method, for the relevant application type(s) (i.e., foliar spray, soil treatment, seed treatment, or trunk spray), please discuss why and what data may be needed to address the exposure route quantitatively.*

Honey and other bees may be exposed to pesticides dissolved in contaminated water via direct consumption and or during contact when worker bees carry water in their "honey stomachs" to the hive for use in hive evaporative cooling. Further, since the latter is done by spreading water onto wax and other surfaces, pesticides not absorbed by the worker bees during water transport may accumulate in the hive and serve as a source of exposure for the entire colony. The Panel notes that water exposure assessments in the white paper are limited to estimates of drinking water consumption, approximately 0.045 mL day^{-1}. The Panel agrees that this volume should be increased to include other water contact. As noted in the white paper this could increase volumes used in exposure assessment calculations to between 0.45 and 1.8 mL day^{-1}. The Panel believes that even when it is assumed that the highest volume is used, and it is assumed that worker bees absorb 100 % of the pesticide transported in water in honey sacks, potential exposures when water from contaminated ponds are used as a water source by bees appear small when compared to dietary or contact exposure described for Tier 1. This is in agreement with findings described in the white paper. The relatively low exposure rate in this case was linked to the relatively low estimated pesticide concentration in ponds using the GENEEC (*GEN*eric *E*stimated *E*nvironmental *C*oncentration) model (US EPA, 2001). In addition, panelists acknowledge that the model generally provides conservative (high) concentration estimates.

In the case of the potential use of puddles in farm fields as bee water sources, the white paper also indicates potential exposures were likely very small when compared to diet and direct contact exposure Tier I estimates. Regarding this point, the Panel disagrees as they note that incorrect units are used in Equation 2-1 found at the top of page 210 in Appendix 2 of the white paper. The Panel notes that the equation is written correctly, but putting values into the equation becomes problematic if all parameters are not in proper units. For application in the exercise in Appendix 2, the Panel recommends the following modified equation:

Here c_w is in units of µg/L, m_{ai}' in units of kg ha^{-1}, d_w and d_{sed} are in units of m, θ_{sed} (porosity) unitless, ρ_b in kg m^{-3}, and K_d in L/kg^{-1}. Re-computation of the worst case puddle scenario used in the white paper (1 lb $acre^{-1}$application to a puddle 0.013 m deep with equilibrium to a depth of 0.0013 cm in the sediment and a pesticide K_d=0.1) yields a concentration of 8100 µg/L^{-1}. The value reported in the white paper is 5.7 ug/L^{-1}. Using the corrected value in calculations indicated that worst case exposure via direct water ingestion is in the 0.4 µg/bee per day. If total water exposure is considered and 100% pesticide absorption is assumed then exposure would be 4 to 15 ug/bee/day. This is within the range estimated for direct contact and diet in Tier 1. Thus, the Panel recommends that water exposure be considered in Tier 1 assessments. Finally one panelist notes that the Agency may want to consider an alternative approach to computing pesticide concentrations in puddles that was developed for ECOFRAM simulations (US EPA, 2012).

The Panel agrees that guttation water expressed on leaves of some plants is a potential bee water source. In addition, the Panel acknowledges that some publications have indicated that pesticide concentrations in guttation water may be high when some systemic insecticides are applied to seeds. The summary of some of these data in Table 2-9 (page 215) is noted with a recommendation that if values are scaled to a 1 lb $acre^{-1}$ nominal rate that maximum projected concentrations should not exceed the water solubility of the active ingredients. Many of the adjusted values in this Table exceed the water solubility of the pesticides identified. The Panel also suggests that guttation water may not be an important water source for bees since the drops are usually found in the morning only. In summary, the Panel agrees that the extent to which bees may be exposed to pesticides in guttation water is unknown.

d. *Other Routes:* Please identify and discuss additional exposure routes (if there are any besides contact with dust and consumption of drinking water) that would contribute significantly to pesticide exposure of bees and explain how and why such exposures could be considered quantitatively in establishing the Tier I exposure value.

A described earlier, a possible route of exposure to some non-*Apis* species is through contact with soil. Many solitary bees nest in the ground, and some cavity nesting bees use mud to partition the cells in their nests. The main route of exposure via soil residues, in these cases, would be topical. Since EPA already has models for estimating pesticide residue concentrations in soils, exposure estimates via this route appear feasible.

8.Tier I Effects Assessments. As discussed in the Problem Formulation (**Section 2.2.1**), the assessment endpoints for the ecological risk assessment of bees involve maintaining honey bee population size, stability of managed bees, quality and quantity of hive products, species richness and abundance. In order to use the results of toxicity studies quantitatively in risk assessment, it is important to identify specific endpoints which will be measured in toxicity tests as these measurement endpoints must have clear linkages to assessment endpoints. As indicated in **Table 1** of **Section 2.2.1**, at the individual bee level (which is the focus of the Tier I assessment), measurement endpoints relevant to these assessment endpoints include: individual survival, adult bee longevity, brood size, brood success, and queen fecundity. The acute and chronic toxicity tests with larvae and adults can be used to quantify effects of pesticides on all of these endpoints, with the exception of queen fecundity (which would require an egg laying study involving the queen). The focus of the chronic toxicity tests with larvae and adults is on mortality that may occur during the tests. Potential impacts of a pesticide on brood size and success can be assessed by determining whether there is decrease in the number of brood (*i.e.*, larvae) following a chronic exposure of larvae to that pesticide. Potential impacts of a pesticide on adult survival and longevity can be assessed by determining the mortality and the decrease in the life spans of adult bees following chronic exposures to the pesticide. The notable limitation to the proposed chronic toxicity endpoints is that they do not include measures of queen fecundity.

a. Please comment on the extent to which currently available bee toxicity tests, which focus primarily on mortality/survival, serve as an effective Tier 1 screen.

The Panel believes that the Agency's articulation of the Tier I bioassays in the white paper is thorough and detailed. The Panel also notes that the incorporation of larval and other bioassays in the Tier I screening process is a welcomed step. The discussions below are aimed to further elucidate how social insects such as honey bees introduce some additional considerations for such screening as a surrogate for all bee pollinators.

These bioassays use the individual as the basal denominator. This does not take into account subtle variation among individual honey bees. Standardizing to honey bee weight would seem to make sense, which might perhaps also translate more effectively to other honey bee castes (queens and drones, which weigh ~1.5x than worker bees) and particularly other bee species, particularly solitary species which are often quite smaller than honey bee workers.

Moreover, it can be quite misleading to test an obligatory social insect, such as honey bees, in an individual context. Thus, it is much more biologically relevant to test honey bees in a social context, ranging from small groups *in vitro* (see below) to large colonies *in vivo* (see Tier II).

The Tier I screens on adults as outlined in the white paper (both topical and oral exposure) also do not capture temporal variation but rather focus on final mortality after a pre-determined period of time (typically 48-hours, but sometimes extended to 96-hours). Standard survival analyses would be much more informative, as it would capture not only the percent mortality at each time point measured, but it would also statistically compare the slopes of the mortality curves. In doing so, quantifying survival in a social environment *in vitro* is much more meaningful than individually for social bees such as honey bees.

Employing two types of Tier I honey bee cage studies would each be informative in different ways: one-time application vs. continuous exposure by feeding. The one-time application bioassay would involve topically treating "bee cups" with a cohort of marked honey bees in a caged group (~50 newly emerged adult workers) and monitoring their survival over time until all focal bees die (which can take many weeks), replacing all dead bees daily with newly emerged workers to maintain adequate group size (e.g., Evans et al. 2009). The continuous exposure application would be to similarly monitor a marked cohort of untreated workers in a cage fed with treated food. Because a one-time dose would not adequately represent bee exposures, chronic tests would better represent real exposures. Also, chronic tests would better represent solitary bee larval exposures because the entire food provision for each larva is usually collected over a relatively short time period (within a day).

Summer and winter honey bees are physiologically different. The timing of the proposed compound and when honey bees are exposed ought to also take this into consideration.

Finally, we need to be cognizant of genetic differences among populations as well as pay attention to behavioral (sub-lethal) abnormalities. As no trigger values are established for sub-lethal effects, the risk assessor needs to have the flexibility to ask for additional data or higher tier tests when concern is raised on the basis of sub-lethal effects in laboratory. While recognized insecticidal substances will always cause sub-lethal effects before mortality, some substances that are not classified as mortally toxic to bees do cause reproducible behavioral abnormalities. As such, the tiered system needs to allow the flexibility to ask for additional data and must not serve as a cut-off value approach.

b. Please comment on additional measurement endpoints (e.g., growth) which should be considered in future modifications of Tier 1 test protocols and which are appropriately linked to the proposed assessment endpoints. Given that the queen is the reproductive unit of the colony, please comment on methods to evaluate effects on individual queens, considering practical limitations of testing with queens.

It would be helpful to include a temporal component to Tier I bioassays, using survival statistics to quantify slopes as well as final percent mortality (see Question 8a). These measurements would then extend well beyond the 48-hour or even 7-day measurement period, and measure the entire lifespan of individual honey bees (often several weeks). While longer in duration, the Panel believes that the increased information gleaned from such comparisons more than compensates for their inclusion in such bioassays.

The question specifically asks for recommendations for effects on queens. Due to reports on queen failure, we need to try to measure effects on queens, even though methods are not fully developed yet. As no defined methods for evaluation of effects on individual queens are available, higher tier tests in semi-field and field settings can be used to evaluate the effects on queen fecundity and also on queen rearing. Effects on queens can be quantified in multiple ways, including effects of rearing environment on morphology and insemination/mating success. Quantifying queen reproductive potential at the individual level can be accomplished in multiple ways (ovariole number, fat bodies, other morphometrics, pheromone production; see Delaney et al. 2011) and can translate to colony-level phenotype, but they do not serve as exact proxies. It

is therefore difficult to distinguish between tiers I and II with respect to reproductives because of their inherent need of a social-rearing environment. Nonetheless, adding queens to the proposed "bee cup" bioassays is possible and simply examine mortality/longevity. Measuring the other sexual reproductives in a colony, the drones (e.g., sperm viability), is important to help fill a data gap in the scientific literature. This is an important measure because drones serve a vital role in the reproductive biology of honey bees.

Quantifying immuno-competence on adults (and larvae) would also be helpful to assess the possible long-term sub-lethal consequences of exposure, but so little is known about the linkages of the immune system to colony health at this point to make such inferences very strong.

Other measurements beside survival may be motility, social interactions, and other behavioral changes. As discussed in Question 8a, sub-lethal behavioral changes at the individual level may have profound effects on the health and productivity of a honey bee colony colony (e.g., Wu et al. 2011). Technology that helps to automate the recording of such data would streamline the logistics of these measurements.

The Proboscis Extension Reflex (PER) and learning is another bioassay and outcome that can be informative. The Proboscis Extension Reflex is a well-studied and common bioassay for learning and memory (e.g., Ciarlo et al., 2012), which uses the instinctive reflex of a harnessed adult worker honey bee sticking out her proboscis when her antenna contacts a droplet of sucrose solution. This reflex can be conditioned to other stimuli, such as odors or tactile cues, so that they associate the two. This enables the proboscis extension as a reliable signal of signal detection and learning. Differential effects of learning, as reported by PER, can distinguish treated vs. untreated cohorts of workers. PER still needs to be linked to colony life/survival, but nonetheless could be used to trigger higher tiered tests.

Another potential assay that could be explored for development could be one that tests the effects of pesticides on queen pheromone production. For example, there is some recent research on how disease impacts pheromone production (e.g., Alaux et al., 2011). Thus, it seems conceivable that pesticides may also have an effect on pheromone production.

The viability of sperm within queens could also be a potentially useful measurement. Also, the fecal samples of queens could be extracted without sacrificing them. Analysis could be performed on the fecal material that may produce some valuable information.

For *in vitro* tests, it is presently uncertain if growth—as measured by weight of larvae—is a reliable factor, but it can be measured on the larvae gained by artificial rearing in the *in vitro* method of Aupinel. Also, the growth of individual larvae can be assessed in higher tier tests by weighing individual larvae, e.g. in the Oomen method (which is a somewhat intermediate between a tiers) or the higher tier test method OECD 75 (Oomen et al., 1992; OECD, 2007). For OECD 213/214 (oral and contact toxicity) tests in the laboratory, the slope should be provided (OECD a & b , 1998).

9 Tier 1 Larval Toxicity Testing. Section 4.1 of the white paper discusses new data requirements for the screening-level effects assessment and recommends obtaining and using larval toxicity data on individual bees. The paper specifically identifies the assay initially proposed by Aupinel *et al*. (2007) as one methodology for quantifying acute oral larval toxicity in the Tier I screen. These assays rely on feeding bees a sugar solution which has been spiked with the test material; however, this *in vitro* method of feeding larvae differs from the process by which the larvae would typically be fed within the colony environment, *i.e*. by nurse bees secreting either brood food or royal jelly.

a. Please comment on the extent to which the Aupinel et al. (2007) in-vitro method serves as an appropriately conservative estimate of Tier 1 acute oral exposure of honey bee larvae to pesticides, given differences in this test design from actual in-hive exposure conditions (e.g., during the first 3 days of the larval development stage larvae consume royal jelly and brood food) and the uncertainty regarding the extent to which larvae rely exclusively on pollen/nectar as opposed to royal jelly/brood food.

Determining the toxicity of pesticides on larvae using *in vitro* testing is complicated due to several factors including possible genetic differences among colonies and difficulties synchronizing larval age, both of which were noted by Aupinel et al. (2007) and mentioned in the white paper. In addition, as yet uncharacterized attributes of royal jelly/brood food, honey and beebread may be relevant to pesticide tolerance in larvae. Royal jelly, for example contains substances that inhibit DNA methylation and histone deacetylase activity (Spannhoff et al. 2011) and promote gene expression. Similarly, honey and beebread contain phytochemicals that upregulate detoxification genes in worker bees (Johnson et al. 2012). These factors have the potential of altering responses to pesticides. Without a more complete understanding of how the composition of food affects toxicity or bioavailability of spiked compounds, conclusively stating that the method is conservative is difficult, although available information points in that direction.

One specific concern is that the proposed method for testing larvae is based on pesticide levels in pollen/nectar administered to 5 day old larvae (p59 of the Agency's white paper). If 5 day old larvae destined to be queens eat royal jelly, and if royal jelly has 100x lower pesticide levels than nectar/pollen, then tests conducted in Tier 1 will have maximum concentrations of pesticides based on worker exposure that are much higher than those experienced by queen larvae. In terms of toxicity this approach seems reasonable, but if caste determination and viability are concerns then pesticide amounts may be inappropriate. As noted earlier, the assumption that royal jelly concentrations are $1/100^{th}$ those found in other foods is apparently based on only one published, Davis and Shuel 1988, and one unpublished, Kamel et al. , study.

The impact of larval diet on pesticide detoxification represents a significant data gap. In the artificial diet used for this bioassay, glucose and fructose serve as proxies for honey; chemically, however, honey is considerably more than just sugar. Phytochemicals in honey may well influence pesticide toxicity and metabolism (as they do in other herbivorous insects*) and interpreting toxicity data in their absence presents a challenge. (Riskallah et al. 1986; Kennedy et al. 1987; Abd-Elghafar et al. 1990; Robertson et al. 1990; Liang et al. 2007)

More generally, the concept of an acute dose for a larva may be biologically difficult to interpret. Unlike workers, which may experience one-time encounters through pesticide drift outside the hive larvae are continuously provided with food that for the most part has been processed by workers, and seem less likely to have such encounters (although they are occasionally fed raw pollen and by that route could conceivably consume an acute dose). Moreover, because larvae are confined during their development, they are subject to continuous contact with wax, which if contaminated is likely to be a source of toxicity not directly assessed by proposed methods. Contact bioassays using pupae might be optimal for assessing topical toxicity.

The *in vitro* method is described and some validation testing has been conducted, but it is currently not validated for the period from pupation to eclosion. As designed, current test protocols would fail to detect effects of IGR (insect growth regulators) on bee brood that are manifested post-pupation. Thus, if pesticides to be evaluated are known or suspected to be growth inhibitors, higher tier tests such as the OECD 75 are needed (OECD, 2007).

The Aupinel test has a high degree of uncertainty associated with variable reproducibility and some additional limitations. It is expected that further discussions (starting October 2012 in the context of the OECD proposal for guideline development) will improve the system. There is also some uncertainty regarding the dosing. For example, there is ambiguity in identifying ecologically appropriate doses in the absence of unambiguous nectar/pollen data.

b. Please comment on the extent to which pesticides would be more or less bioavailable using the synthetic matrix relied on for feeding developing bees in this in vitro method.

Bioavailability will be influenced by the nature of the compound, larval metabolism, and food composition. Without accurate information about how these factors influence bioavailability within honey bee larvae, it is not possible for the Panel to assess this question. Aupinel et al. used dimethoate, an organophosphate that is highly water-soluble; it is not clear how pesticides with different structures and properties will perform in the synthetic matrix.

c. Please comment on the extent to which the absence of trophallaxis (i.e., the transfer of food/fluids between colony members) may render larvae more or less vulnerable to pesticides.

The Panel believes that contributions of trophallaxis to pesticide resistance or susceptibility are, to our knowledge, totally unknown. This lack of knowledge in the research represents a data chasm.

d. Please comment on alternative methods for testing individual larvae that may be appropriate for quantitative use in a Tier I screening-level assessment.

Literature reports of alternative methods for testing larvae quantitatively are vanishingly rare. Atkins and Kallum 1986 treated larvae in a colony by micropipetting pesticides directly into the cells. (Atkins & Kallum 1986)

Pupal bioassays have not been widely used for pesticide toxicity, but they have been used in other contexts (e.g., pathogenicity of fungi). Furthermore, these studies may be productively employed. Hendriksma et al. (2011) is noteworthy in this regard; this bioassay uses a nongrafting method to manipulate larvae and uses prepupal weight as a bioassay endpoint to calculate LD50 values for dimethoate). A few other protocols for testing effects of pesticides on honey bee larvae are available, (e.g. Wittmann D., 1982). In addition, some methods have been described for larval testing of *Osmia* and other bees but these have not yet been validated or ring-tested.

Larval transcriptomes and methylomes of honey bees have been described. The existence of these data provides potential quantitative measures of effects of toxins on larval development that might be useful as alternative indictors of larval actions/health. Genetic resources may allow use of gene expression, detoxification and/or antioxidant enzyme activities, senescence traits (vitellogenin) as biomarkers. Johnson et al. 2009 suggested the use of ribosomal RNA fragments as a biomarker. Such methods, however, would need validation before they could be used in Tier 1 screening.

e. Typically acute toxicity tests are concluded between 48 – 96 hrs. Please comment on the appropriate duration of toxicity tests for assessing acute toxicity to individual larval and adult bees.

There are problems interpreting larval survival over a 2 to 4 day period as a metric of toxicity. The four-day endpoint represents 80% of the period of active feeding, so distinguishing between acute and chronic becomes difficult. Some agents may induce toxicity due to delayed larval development, which would be overlooked in the proposed approach.
In general, a biological transition, such as pupation, provides a more reliable, consistent, and interpretable endpoint. Furthermore, the endpoints-time to pupation, pupation rate, and pupal weight are more biologically meaningful, and different castes (queens and drone)s can be assessed separately. Such tests, however, have not yet been vetted thoroughly. Assessing acute toxicity over seven days for the present time will allow EPA to harmonize international efforts at the OECD level until a satisfactory standard pupation assay has been developed but developing such a test should be a priority. Among other things, such assays will be more sensitive to pesticide impacts on growth rate and development (e.g., IGRs).

Currently, the Aupinel method describes both single dosing vs. repeated dosing of the substance. The latter (dosing of the food every day) seems to be the most realistic, but there may be a need for two toxicity standards. This is because there may be substances that are acutely toxic to young larvae (uncapped brood) and some that only affect later brood stages (capped brood, e.g. fenoxycarb; Aupinel *et al.*, 2007). More effort should be focused on establishing a validated, and reliable method that also covers the time from pupation to eclosion.

10. Tier 1 Chronic Toxicity Testing Bees. **Section 4.1.2** of the white paper discusses the status of chronic toxicity tests with individual adult and larval bees. At this time, no formal guidelines have been developed for conducting chronic toxicity tests with either adult or larval bees, although studies with individual bees of various ages are routinely reported in open literature.

a. Please comment on the conclusion that adequate procedures have not been sufficiently developed and validated for assessing chronic toxicity to individual bees in a risk assessment context.

The concept of the "individual adult bee" presents a challenge for assessing toxicity in *Apis mellifera* because of the importance of the social context. Age-related polyethism over time is circumvented in the absence of a social context and foragers are hard to keep alive individually. However, exposing emerging workers for ten days at low dose with known ages of bees has been attempted according to OECD 213 with some success (OECD, 1998). Chronic tests in the laboratory can be conducted by using the protocols of OECD 213 (with small adjustments, e.g., using freshly emerged bees). Caged brood tests or broodless "bee cup" tests may be employed productively in this context.

b. Please comment on the potential use of the 10-d adult worker and 7-day in vitro larval toxicity tests discussed in the white paper for assessing chronic toxicity once these methods are fully vetted.

The optimal design for a larval *in vitro* test would be one encompassing the entire active feeding period through pupation, with time to pupation, pupal weight and percent pupation as quantifiable and interpretable measures of pesticide impact (As demonstrated by Hendriksma et al. 2011, with prepupal weight and survival to prepupation as assay endpoints. Such a test can provide data on both acute and delayed responses). Obviously however, such testing would require vetting. With respect to adults, use of a 10-day viability test for adult bees has the virtue of allowing for harmonization with ongoing efforts in Europe.

c. Although 10-day adult and 7-day larval toxicity tests have been proposed, please comment on whether alternative durations of pesticide exposure may be more appropriate for determining chronic toxicity for adult and larval bees at a Tier I screen.

A ten-day test of adult survival can be informative, but interpreting the biological meaning outside the social context presents a challenge. Rather than an alternative duration of exposure, an alternative design should be considered. Survival could be assessed in the social context with the use of caged brood or broodless bee cups (although not a standard methodology), which have been used with success. As for larval toxicity tests, as stated earlier, extending the test duration through pupation rather than terminating it at an arbitrary interval would be most appropriate.

d. The white paper identifies NOAEC No observable effect effective concentration as the chronic toxicity measurement endpoint. Please comment on the possible use of EC_x values as a measure of chronic toxicity for use in RQ calculations.

The proposed regression-based method for calculating ECx levels seems feasible if a sufficient number of concentrations can be tested to allow for confident estimates (confidence levels drop at low and high ends of dose-response curves). Including the temporal component of chronic effects provides a more quantitative method for estimating effective concentrations causing

chronic effects (e.g., survivorship curves). Unique attributes of bees favor anything that works well with small sample size and high variability.

NOAEC values may depend on selection of endpoints and appropriate concentrations for assay. Both NOAEC and ECx values should be provided wherever possible. There are some uncertainties and difficulties with linking the outcomes for both methods with the protection goals, but they are still useful for triggering higher tier tests. The European Food Safety Authority opinion (EFSA, 2012) may have useful suggestions as to how chronic tests can inform risk assessment.

e. Please provide comments on what percent effect would be considered a relevant measure of chronic toxicity for individual bees given the potential compensatory effects which honey bee colonies may exhibit relative to the effects of pesticides.

Compensatory effects are a societal phenomena that are better suited for Tier II assessment. Recruitment of precocious foragers provides a short-term solution to colony needs but may exact long-term costs on colony health. Khoury et al. 2011 estimates that reduction in worker longevity by 2.8 days corresponds to a loss of 0.35 bees per day. Incorporating a time component may be useful.

Thus, EC_{20} (20% difference between controls and experiment to trigger concern) may be a conservative point to set, given that, according to the one model available (Khoury et al. 2011), loss of 30% or more foragers daily triggers the death spiral (losses at that level are irrecoverable). Additional modeling can be useful here, inasmuch as this estimate is not based on an abundance of empirical data. Timing of loss would matter, too (losses are more sustainable in the fall season. Thus, the EC20 may vary depending on the season).

f. Although the white paper identifies some measurement endpoints for assessing chronic toxicity to individual bees (e.g., survival), please comment on other endpoints to consider in chronic toxicity studies which the Panel believes are important and the associated study design elements.

As mentioned in the response to Question 8, other endpoints to consider in chronic toxicity tests include behavioral impacts (e.g., learning assays, automated recording of motility and behavioral interactions). Future endpoints could include gene expression relevant to detoxification and/or antioxidant status, immunocompetence, and senescence (e.g., vitellogenin). Johnson et al. (2009) suggested the use of ribosomal RNA fragments as a biomarker. Such surrogate biomarkers (an approach widely used in human risk assessment) need to be validated for bees.

g. ***Section 4.1.3 4.1.2.1.2.*** *of the white paper discusses the uncertainties associated with developing risk assessments based on studies of sublethal effects when sufficient linkages have not been developed to understand how the sublethal endpoints may be quantitatively related to typical assessment endpoints (i.e., growth, impaired survival, and reproduction) at the whole colony level. Please comment on the proposal to use data on sublethal endpoints to qualitatively (i.e., no Risk Quotient is derived) characterize effects and risk until sufficient*

linkages between these endpoints and the defined assessment endpoints have been developed (e.g., Adverse Outcome Pathways).

The social environment complicates the process of identifying reliable, reproducible and relevant sublethal endpoints. Given these complexities, the use of sublethal endpoints is conservative and important for providing insight into super-organism responses. They are also critical triggers for signaling a need for moving to Tier II.

A high priority for future research includes the development of assessments for neurotoxicity symptoms; although direct links to colony health may not yet be fully understood, they are indicative of biological activity of chemicals. The results of these types of studies can demonstrate a need for additional quantitative tests at Tier II. Eventually, links must be made. Proboscis extension reflex, albeit a high throughput assay for behavior, measures appetitive learning but many other forms of learning are essential to colony health and few measures exist for documenting impacts on those forms of learning.

11.Tier II Semi-field Effects Assessments (Whole Colony). For Tier II assessments, **Section 4.2** of the white paper identifies two types of test methods that may be used to assess colony-level effects, *i.e.*, semi-field tunnel tests [OECD 75; EPPO 170]; and semi-field feeding studies (OECD 2007; EPPO, 2001). These studies are intended to help characterize risks identified in the Tier I level assessment that are based on exposures and toxicity data for individual bees and quantified using Risk Quotients.

a. Although colonies are typically confined to enclosures for Tier II studies and these enclosures can limit the environmental realism of the study conditions, tunnel studies provide an opportunity to collect colony-level effects and potentially exposure information. Please comment on the relative strengths (e.g., foraging activity by adult worker bees is limited to treated crop; trophyllaxis enabled) and limitations (e.g., limited study duration, smaller colony sizes, reduced forage area) of these methods.

Simple LD_{50} measurements of topical toxicity to a solitary adult worker honey bee do not adequately capture the myriad outcomes of exposure for social insects. This is because there are other deleterious sub-lethal effects beyond simple adult mortality due to exposure. Thus while quantifying contact toxicity by contact and ingestion is an important dimension to quantify, it is only one side of a multi-faceted problem. This is a fundamental issue that faces the Panel. Because honey bees are social insects and function like super-organisms, the colony is the basal unit to test toxicity. As such, *in vitro* approaches are limited because they do not test at that higher level of biological organization. The analogy for mammalian systems, for example, is to screen compounds on tissue cultures and assume an LD_{50} on the whole animal from cellular information. While informative, it does not adequately capture the entire biological reality of the system in question. Employing the additional bioassays discussed in Question 8 would help bridge this gap and serve as effective screens, but they will still not fully serve as an adequate proxy for testing colony-level effects.

Strengths of the proposed Tier II studies in the white paper are that the effects are measured *in vivo*, rather than extrapolated to a complex environment from a simplistic *in vitro* bioassay.

Because of the additional level of biological complexity, the Tier II level for a super-organism is the basal level in which to properly quantify LD_{50} estimates.

Limitations to this approach are increased among-colony variation, which requires high samples sizes to detect weak but biologically relevant effects, as well as ability to standardize across different trials. Because it is difficult to acquire a sufficient sample size due to the inherent variability between colonies, there are increased risks of Type II errors (false negatives) with high-tiered tests. Modeling may help facilitate data collection and analysis. The Panel also notes that that honey bees can also buffer potential effects because of their sociality. Consequently, this may obfuscate effects that may be important for solitary bees.

The semi-field design is a worst-case enclosure appropriate for determination of acute effects on bees. The exposure is maximized in the tunnels/tents as bees are confined to the treated area; where dose responses (by application of different field rates) and mitigation measures (e.g., application after bee flight instead of application during full bee flight activity) can be tested. Thus, with semi-field tests a very high degree of certainty is observed. One further limitation is that some colonies respond with a reduction of the breeding/larval rearing activity, making interpretation of effects on colonies under real conditions more difficult. For addressing the individual larvae, the OECD 75 may be used (OECD, 2007).

Standardizing by weights of individual bees, as well as for colony size, will be helpful to standardize across colonies, studies, and other systems. In doing so, there is a need to determine linear vs. non-linear effects based on exposure concentration.

Also bumble bees can be used in Tier II studies, particularly for caged designs. The possibility of using *Bombus* as a model system to test social effects may be more effective in some ways because they are amenable to field-realistic foraging behavior in confined spaces. It would be helpful to test both *Apis* and *Bombus*, but if *Bombus* is used for semi-field analyses it would help translate between them. While both may be utilized, *Apis* may be more amenable to feeding experimental designs, whereas *Bombus* may be more amenable to cage designs.

It is a general question as to how feeding methods fit into Tier II tests. For example, the Oomen test is regarded as an intermediate between Tier I and Tier II, but in general the feeding tests are not acceptable as Tier II tests especially under free flying conditions (Oomen et al., 1992). In the case of a combination of a feeding method with colonies free-flying, a minority of the panel members believe that the studies should be rejected and not considered as valid as the exposure cannot be quantified. Exposure can only be quantified and assessed if these tests are conducted in semi-field conditions in closed tunnels; the bees must not have access to other uncontaminated food sources. Other panel members, however, feel that while precise exposure levels cannot be obtained, they can still be estimated and quantified (amount of food consumed over time, quantifying alternative foraging activity) and therefore have some utility. Enclosures limit the time a test can be run because it is very difficult to provide sufficient bloom resources (pollen and nectar from flowers) in an enclosure. Thus, enclosures might not adequately represent field exposure levels (if exposures are based on residues on treated crops within the enclosure), thus a better method might be free-flying but fed pesticides.

Feeding with spiked food should only be used to extend exposure where necessary, and it is not possible to do a design with extended flowering of the realistic crop (e.g., if there is concern for chronic long-term exposure). For some crops (e.g., maize), it may be possible to extend the duration of exposure also by planting crops at different time intervals, but for other crops (e.g., winter oilseed rape) this may not be possible. In these cases, the treated crop of concern should be used for realistic exposure during flowering and thereafter could be extended with feeding both nectar/sugar solution and pollen/pollen patties enriched with field realistic concentrations. In all feeding studies, it should be kept in mind that only both contaminated nectar and contaminated pollen are fed. Exceptions in terms of using only spiked nectar/sugar solution or spiked pollen are only acceptable if a special use of a special crop is investigated (e.g., maize, which only supplies pollen).

b. Please comment on any other types of colony-level studies that should be considered as part of Tier II.

Alternative study designs that may be further explored include observation hives to record behavior and estimate interactions. This directly addresses questions about sub-lethal behavioral effects outlined in Question 8a. Moreover, it would be very straight forward and highly informative to rear queens and drones and to see the effects on reproductives. This could be done in either semi-field experimental designs (enclosed flight cage of free-flying colonies fed supplementally). Finally, the Panel believes that there may be *in vitro*, whole mini-colony tests that could be implemented in the short term that blur the lines between tiers I and II. For example, Tier I tests of "bee cups" each containing a queen could be verified with Tier II tests of "baby nuc" colonies housed in incubators. While still *in vitro*, such colonies would contain queens, brood, comb, and other full-colony phenotypes that could be measured.

With all tests that include artificial feeding, as an intermediate test between Tier I and Tier II, the Oomen Test can be used (Oomen et al., 1992). Limitations are identified so the original method should be optimized to ensure exposure (e.g., the feeding period should be extended to feeding contaminated sugar solution not only on one day but for more than 9 days). Also, this method has some pros and cons, some of which are also described in the EFSA opinion (EFSA, 2012). For other feeding methods, there are no guidelines available. Historically, only the Oomen method has been used.

Some special study designs can be used for investigation of special effects or exposure routes. Examples include studies to assess the effect of a pesticide on queen rearing, (e.g. a shook swarm") or "swarm box" method of queen rearing, exposure via the crop and observation of the number of reared queens and rearing success of hatched queens. This may be an important area of risk assessment since studies have shown that queens are highly sensitive to some pesticides during development and beekeepers report anecdotally that queen longevity appears to be reduced in the past decade or so.

Another example is investigating the potential effects of guttation. In semi-field studies, controlled conditions in tents offer the possibility to simulate water collection from guttation droplets and other water sources. Clean water sources can be excluded to ensure maximum exposure. The effects on foragers and hive bees and different brood stages can be measured in

worst-case exposure scenarios. Nevertheless, it is pointed out that semi-field studies may have a limited potential for extrapolation of the findings to field conditions.

c. *Please comment on the most important endpoints that should be measured in the Tier II studies (e.g., adult forage bee mortality, brood development, queen fecundity, overall colony strength) that are linked to assessment endpoints and their associated protection goals.*

Measuring different colony phenotypes is important. Calculating change in each, as well as using Principle Component Analyses (PCA) for generating a composite variable, would be most effective. Measuring queen and drone viability would also be helpful.

The Panel also sees a need to verify foraging behavior on the crop to quantify exposure rates of the foragers. In other words, there needs to be a documented association between foraging activity and colony phenotype.

Tunnel studies can vary by experimental design depending on the expected exposure route, either with the target crop in the tunnel or not. Using a crop results in uncertainties about dose, but panel members disagree about their relative utility. If the exposure is expected to be primarily through pollen collected by bees, then a spiked pollen patty is appropriate. If it is expected to be through contact with the crop, having a treated target crop is appropriate. Spiked syrup can be used if there is concern for exposure to systemics via nectar.

Tunnels should be air conditioned and sufficiently large to keep the maximum temperature at 85°F or less. It works best if the nuclear hives (nucs) are moved in with mostly bees that had not foraged. A water source such as a tray with gravel and water should be provided.

For all tests with bee colonies, the behavior at the hive entrance as well as the behavior of bees which are foraging on crop should be assessed. There is a need to sample and process returning foragers to see what they are foraging on.

Queen fecundity may best be addressed by the assessment of the amount of brood in the hives at certain intervals, the effects on larval development and successful brood rearing with the OECD 75 method, using maps or digital imaging of the brood (OECD, 2007).

d. ***Section 4.2.2*** *of the white paper discusses a full-field feeding design. This methodology is discussed under Tier II assessments since the colony is relatively confined to foraging on either spiked sucrose solutions or spikes pollen. The intent of this methodology is to ensure that colonies are exposed to known residue levels over longer durations than the semi-field tunnel study designs. A limitation to the study is that bees may simply store spiked food rather than consume it and that the reliance on a single source of food may introduce confounding effects (e.g., nutritional deficits) into the study. Please comment on the environmental realism and utility of full-field feeding studies as a line of evidence in characterizing risk to honey bee colonies.*

Direct feeding is helpful, as it gauges known consumption. However, it detracts from realistic field effects and contains inherent aspects that are not altogether meaningful. It may be more helpful to sample incoming foragers, processing pollen and nectar loads from each in order to determine the variation and degree of exposure. It is also necessary to control for the amount consumed.

With increased variation, which is always the case with colony-level phenotype in an ecological setting, replication and large sample size is paramount. Forced feeding eliminates the natural foraging behaviors, including colony communications and avoidance behaviors.

Twenty-eight days may not be long enough to detect certain sub-lethal effects. Two brood cycles (42 days) would quantify not only affected foragers but also the bees raised in the controlled environment. Even if left in tunnels for a shorter period of time, extending observations afterwards seems critical.

With the proposed feeding tests, an underestimation of exposure is possible or even very likely. Free flying is not considered appropriate by some panel members as additional uncontaminated nectar and pollen will be foraged, thus there is no worst case exposure. Feeding in enclosed semi-fields is considered adequate but feeding studies should be designed carefully to ensure that no underestimation of exposure is possible.

Full-field feeding studies are not considered appropriate for risk assessment purposes. At best, they can be used as supplemental information of limited value as they are not environmentally realistic.

In general, it is recommended the methods as described in OECD 75 and EPPO 170 be used for Tier II and Tier III studies (OECD, 2007; EPPO, 2007). As there may be effects by contact and/or oral exposure, the majority of the Panel feels that these study designs are appropriate to test the exposure to both pathways and the combination of both as it may occur in realistic conditions.

If carefully designed, feeding studies in general only cover oral but not contact concerns (see above).

e. *As discussed in **Section 4.3.4** of the white paper, it is important to consider the biological significance of a measured effect in addition to its statistical significance. Please comment on the nature and magnitude of effects that would be sufficient to conclude biologically significant effects on the colony and/or the need to transition to Tier III assessments.*

Power analyses would be informative on this point. If sample size is small, a power analysis would indicate the biological significance of the results. Replication and meta-analysis would do the same. Effects on fertility should trigger Tier III testing, or the ratio of mortality to reproductive potential (small effects on reproduction are more important if coupled with a high bee mortality from toxicity).

Ground-truthing with transect walks of foraging bees would quantify, at least to a certain extent, the foraging capacity at the target crop. Sampling returning foragers would also be very informative, although they would require greater inputs to obtain those data.

The Panel notes that the step-wise process that assumes a low risk *in vitro* immediately equates with a low risk *in vivo*. Because of unforeseen synergistic interactions, sequestration, and concentration, Tier II and III studies should be performed regardless as determined by the risk assessor, but the degree and sample size may vary depending on the findings at lower tiers and the likelihood of pollinators encountering the pesticide in its application.

Combining all similar studies in a common database would facilitate further inferences by quantifying expected colony phenotypes.

Single metrics can be misleading in their independent importance (e.g., worker mortality in absence of an understanding that workforce strength is the ongoing demographic process involving the ratio of death to eclosion). Nonetheless, the Panel has attempted to prioritize the list of potential measurement outcomes listed in the white paper on pages 132-133 (see Table A) as a means to orient risk assessors towards more-informative measurements that have greater penetration (i.e., they likely measure a stronger biological signal in the system).

Table A. Prioritized potential measurement outcomes

Relative utility for measuring risk	Ecological measures	Colony measures	Individual measures
Higher utility	Mortality in the crop	Colony strength	Pesticide residues
	Mortality at the hive	Weight of the hive	Longevity
	Mortality of drones and pupae	Food stores	Behavioral abnormalities
	Foraging activity in the crop	Presence of the same queen	
	Returning foragers	Ability to requeen	
Lower utility		Levels of disease Over-wintering success	

It should be noted that "colony strength" is a general term and is a non-empirical composite of many inter-correlated factors (e.g., adult worker population, brood area). Ideally, many of these measurements can be made and then compared using Principle Component Analysis to relate overall colony phenotype (i.e., colony strength). The potential endpoints listed are useful and appropriate to look for potential adverse effects at the colony level. The endpoint listed as "the returning of foraging bees" should be listed as "colony foraging activity measured as bees returning / 3 minute period." Also, "weight of hive" should be listed as "change in colony weight." Lastly, a measure of the consistency of the "brood pattern" can be made by counting the number of open cells in several (3-4) 100 cell areas across 3-4 frames containing sealed

brood. Poor brood patterns, while perhaps not indicative of pesticide exposure, have been linked to poor colony performance and or increased likelihood of colony mortality.

Expert judgment is needed to interpret the statistics with respect to biological significance. Depending on the statistical power and certainty needed, statistics are possible but need very careful understanding. Absence or presence of statistical significance is not an indication of no effects or acceptable effects or respectively of adverse effects. The different endpoints may have different statistical value for the interpretation.

12. Tier III Effects Assessments. **Section 4.3** of the white paper discusses the proposed risk assessment process in Tier III that relies on assessing effects at the colony level where colonies are not confined (*i.e*., full field) and exposure is intended to represent environmentally realistic conditions. As discussed in the white paper, interpretation of the biological significance of colony-level effects can be challenging, regardless of their statistical significance. Conversely, high variability in field studies can limit the statistical power of the study to detect treatment effects. The paper identifies uncertainties associated with the extent to which bees forage on the treated crop, the size of treatment site, controlling for alternative forage/pesticide use area, and ensuring suitable controls; these factors combine to render these studies highly resource intensive.

a. Please comment on the strengths and limitations of full field studies described in the white paper.

Tier III studies should be designed to provide real world exposures and integrate their potential effects at the colony level. However, these studies have many potential pitfalls with regard to their ability to carry out a controlled experiment in the open environment with adequate controls. Such tests are already widely used in the EU. When properly executed, Tier III studies are the best means to assess the risks and or impacts of pesticides to *Apis mellifera* colonies and other managed pollinators e.g. *Bombus* or *Osmia*. In the context of the regulatory framework Tier III studies are considered "highly refined" studies meant to address specific concerns raised in lower tier studies. Thus, each study will vary depending on their measured endpoints to address specific concerns raised in Tier I or II testing. There are many strengths and weaknesses in doing Tier III testing. The Panel identifies them below:

Tier III study strengths:

- The natural exposure that can be achieved by allowing bees to forage freely on a treated crop, giving the test more relevance to real-world situations.
- The measurement of endpoints to test for effects at the colony level that are unattainable at the Tier I and II level testing.
- The exposure achieved can accurately simulate the actual use of the pesticide.
- Tier III tests allow for colony level testing of protection goals. The use of full size colonies will provide information on colony development, multiple exposure routes and additional, non-pesticide related stressors, such as factors and endpoints which can only be assessed to a limited extent in tier II tests.

- If significant effects are noted at the colony level then the results can be interpreted with a high level of confidence as the colony has many buffering mechanisms such as a reserve worker force, ability to compensate or replace dead or impaired worker bees etc.
- The use of surrogate crops such as *Phacelia* is possible to help increase foraging on the crop and thus exposure.

Tier III study weaknesses:

- Increased variability among colonies, and thus there is a significant need for replication and large sample size.
- Study site selection is critical as competing crops or vegetation can result in low to no exposure to the target crop, especially if the plot size is limited or the crop is marginally attractive.
- Plot size will vary with the less attractive crops needing larger plots to insure adequate foraging and thus exposure to the target crop.
- Bees will forage on the most attractive plant within flight range and this can lead to inadequate exposure and high variability in results.
- Pests may need to be controlled in the control plots, and thus may need to receive pesticides that could compound effects of the pesticide being tested. Positive control plots can be expensive, and thus people will want to avoid them.
- Studies with sufficient levels of replication to provide appropriate sample power in varying climatic regimes and eco-regions where target crops are produced may be very costly.

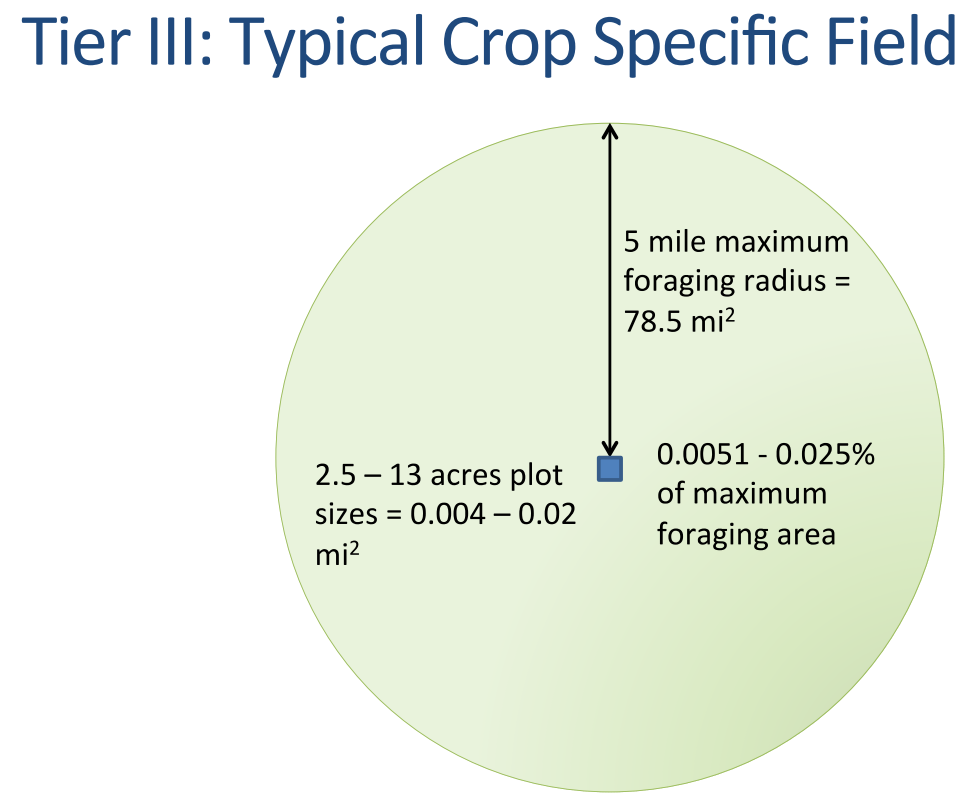
In regard to evaluating the potential interactions of multiple pesticides, complete exclusion of this evaluation runs counter to the goal of Tier III assessments being real-world and biologically relevant. While it would be infeasible to test such interactions in a pairwise matrix, using known and commonly exposed pesticides as a background environment seems paramount.

Consolidating data into a common database with repeated meta-analysis would help in evaluating varied information. Making these data public would help transparency and reassurance.

*b. Please comment on the proposed modifications to the field study design elements presented in **Section 4.3.2** of the white paper.*

One of the concerns raised by the Panel regarding Tier III study design is a lack of control regarding bee foraging. This leads to uncertainty regarding pesticide exposure. This lack of control is implicit in a "field" study. However, there may be ways to estimate the foraging activity and further ways to reduce the concerns regarding estimation of pesticide exposure. For example, as illustrated by EPA during the public meeting the maximum foraging radius for a honey bee is 5 mi. If a field study had plot sizes of greater than or equal to the 78.5 mi^2 area of this circle the probability of bee foraging on treated crops would be 100%. Furthermore, the pesticide exposure can be quantified by analyzing a relatively small number of samples of foraging bees over the duration of the study. Of course, this size field plot is impractical and virtually impossible to implement in a field study. However, this does represent one end

member of an infinite set of circles with decreasing radii from a maximum of F to the origin or 0 mi. Furthermore, if we assume that the likelihood of flying distance x from the origin can be described by an exponential decay curve, e^{-x}, we can develop an equation to describe the expected use, P, of a given plot by bees. The measure of expected use, P_1, by bees of a circular plot with a given radius x where x>5 mi can be described by the following equation:



$$P_1 = \int_0^z \omega_1 * 2\pi e^{-x}\, dx$$

where P_1 is the measure of expected use by bees for plot 1 with radius x, $\omega_1$represents the attractiveness (both quality and density of forage) over area of circle with radius z about the origin.

This equation can be rearranged to the following:

$$P_1 = \omega_1 * 2\pi \int_0^z e^{-x}\, dx$$

Integrating from 0 to z gives the following equation:

$$P_1 = \omega_1 * 2\pi * (-e^{-z} + e^{0})$$

Following the same rationale, the likelihood of bee use of the annulus surrounding the field of radius z and extending to the maximum radius F can be described by the following equation:

$$P_2 = \omega_2 * 2\pi(-e^{-F} + e^{-z})$$

where P_2 is the expected use by bees for the annulus surrounding the field plot with radius extending from z to F, ω_2 represents the attractiveness (both quality and density of forage) over area of the annulus with minimum radius z and maximum radius F about the origin.

Finally, assuming bees can fly anywhere inside the maximum circle with radius F, the probability of plot usage, A, can be quantified through the equation:

$$A = \frac{P_1}{P_1 + P_2}$$

If one assumes that uncertainty in the exposure term is inversely proportional to A, strategy for foraging bee sampling to quantify pesticide residues can be designed appropriately (e.g., frequency of sampling increases as A decreases).

It should be noted that one assumption of this exercise is the relationship used to describe bee foraging. The sensitivity of this assumption can be assessed by changing the form of the relationship (e.g., exponential, linear, logarithmic) or changing the type of exponential (e.g., e^{-x}, $e^{-0.5x}$, e^{-2x}).

Regardless of how risk assessment accounts for differences in foraging area, incident data show that carefully designed Tier III studies are well suited to assess the risk and thereby protect managed bee colonies. In the EPPO guideline 170, minimum values for the number of foraging bees are described to ensure that sufficient exposure is achieved (EPPO, 2001). As described in the guideline, the minimum number of colonies should be 4 per test site, depending on the available forage. For increased test field size, larger numbers of colonies are appropriate.

The Panel also notes the following about Tier III assessments:

- The Panel agrees that plot location is a critical consideration, and competing vegetation is an important factor to ensure exposure to the target crop. Plot size will be affected by this, with less-attractive crops needing larger plots to insure adequate foraging and thus exposure to the target crop. As such, we believe that study site selection is perhaps the single most important element followed closely by determination of adequate plot size.
- Distance between control and treated plots should be a minimum of 4-6 km but increasing this distance must be balanced against having plots in different microclimatic conditions as the distance between plots increases.
- Testing of formulations at highest label rate is valid, as this provides a “worst case scenario” for a given pesticide. In doing so, the exposure achieved can accurately simulate the actual use of the pesticide. However, limited plot size results in a less than worse case test scenario; it is not until the pesticide is in actual use that large acreages may be sprayed and the full extent of exposure and potential effects can be fully evaluated.

- The Panel agrees that balancing or equalizing bee colonies is an important step to be taken in advance of placing hives on study sites; otherwise, colony variation is too great to see differences in colony growth and other measured endpoints. The Panel concurs that opening of hives can result in increased disturbance and the possible introduction of variation within the results (e.g., increased queen loss) but this should not prevent the inspections needed to collect endpoint measurements. Good beekeeping methods can alleviate this concern, which can be readily accomplished by utilizing skilled and experienced beekeepers.
- Power analysis is a useful tool to determine the number of colonies needed to produce statistically significant results. However, the number of colonies needed may not be feasible due to cost constraints but can be offset by good endpoint measurements and observations on changes that are considered biologically significant (adult bee mortality) but may be statistically not significant due to small sample size.
- The Panel concurs that the duration of a field trial can be variable and will be dictated by the endpoints of interest.

One addition to the traditional field study would be the addition of *Bombus* or *Osmia* to the field as they fly shorter distances and thus the level of exposure can be greater or more controlled and also the colony sizes are smaller and solitary females will be tested under field conditions. Additionally, studies (e.g., Whitehorn et al. 2012) have shown that reproduction can be impacted with sub-lethal exposure at field-realistic exposure levels making these solitary or smaller social colonies more sensitive to pesticide exposure. The use of these additional bee species in a Tier III study would not add significantly to the costs and they overcome some of the problems associated with the long flight range of honey bees.

c. Please comment on factors that should be considered in evaluating the biological significance of effects measured in full-field studies in relation to the proposed assessment endpoints and related protection goals.

All the weight of evidence achieved in the different tiers including information on acute mortality of adults and larvae as well as colony development, behavior and behavior need to be taken into account.

Colony growth patterns, as measured by brood production and or adult bee population, are robust measures of colony health and directly related to the protection goals.

Similarly, colony mortality would be a biologically adverse effect, but direct colony mortality is unlikely to be detected because of other factors, and thus a consideration of other endpoints and their relative importance is warranted.

Effects on queen survival or replacement would be of biological significance and have implications for colony survival that may not be accounted for if test duration is only a few weeks or months.

Similarly, overwintering success is the ultimate endpoint, but the more temporally removed the exposure period is from the winter period the more likely that confounding factors (e.g., parasites and pathogens, foraging conditions) will come into play making interpretation of results difficult.

Adult mortality is a good biological indicator as are foraging rates and adult bee longevity; each of these can be good indicators of adverse effects but the effects could be transitory in nature, and thus extrapolation to colony-level impacts are not necessarily direct. Taken together any effects noted in the above endpoints could be used in the risk assessment decision making even if each endpoint alone was not significant. Colonies are complex super-organisms and thus several endpoint measures may be a more appropriate than relying on any one endpoint alone.

Disease and pests levels, worker longevity, foraging rates and most other endpoints measured at the colony level are less robust than those listed above and should be considered as contributing to a measure of colony success but may not be of biological significance by themselves in relation to protection goals. Adverse effects noted in these less robust endpoint measures should be taken into consideration with earlier results from Tier I and II testing.

As long as the plots are not small, one could sample bee diversity (in addition to honey bee measures). A simple method is colored bowl traps traps, where yellow or blue bowls are filled with water and soap, then placed in the field for a pre-determined duration to trap bees (e.g., Tuell et al., 2009).

d. Please comment on factors and methods that should be considered when extrapolating observed effects at the colony level in semi-field and field studies to those expected to occur in the environment (e.g., spatial and temporal scale of exposure, hive management practices, presence of multiple chemical and non-chemical stressors, etc).

Tier II and III studies cannot incorporate all seasonal and colony management practices that could change the response of colonies to pesticide exposure. If testing is conducted in the normal growing season then an adequate level of confidence can be obtained with the results generated. As has been stated earlier, colonies used in either field or semi-field studies must be uniform in strength and as disease and pest free as possible at the start of the study.

The seasonal nature of colony growth and the importance of colony preparation for winter (extra stores of honey and pollen and the production of winter bees) should be taken into consideration when interpreting colony level results. If crops of interest are to be grown primarily in the late summer then tests should be run in this season to make them relevant to the real world.

The longer a Tier III study continues the more likely that outside variables (e.g., pests and disease, other contamination of hive products) can influence the results. These outside influences can be minimized by moving the colonies to a remote common area, away from urban and agricultural settings.

If a crop is major and broadly planted then, the best field test may still not simulate the level of exposure bees may receive once the pesticide is in wide use (e.g. corn or soybeans in the Midwest).

Other agrochemicals need for successful crop production need to be selected carefully since synergism and or antagonism are concerns that may confound results.

*e. A number of study design elements are discussed in **Section 4.3.5** of the white paper; however, even in the best designed studies, there can be confounding effects which can limit the utility of these studies in risk assessment. Please comment on factors that should be considered in determining the utility of field studies for pesticide risk assessment, including a discussion of the representativeness of a study for a National Level assessment (i.e., the pesticide may be used anywhere in the United States and its territories).*

The design elements of plot size, and plot separation as stated above are critical. However, the measure of exposure to identify incoming pollen during the exposure period is critical to ensuring the study's validity. This is the single measure that gives the highest level of confidence that exposure occurred and, residue analysis of the pollen can address this issue.

The ideal test would use the formulated product on the crop of interest that has the highest level of attraction to bees and is most widely planted. This is because the formula might affect residue levels and attraction to the bees (in addition to possible toxic effects that the formulation might pose). A past example was a microencapsulated pesticide that bees tended to collect like pollen, increasing their exposure over what was anticipated. If visitation to the crop is confirmed and no adverse effects are noted, then use of the product could be assumed safe with a reasonable degree of confidence.

If the pesticide is to be used on a widely planted and common crop in the US, then testing should be representative of the environments in which the crop is normally grown (i.e. tested in several different geographic regions). Additionally, climate change is a continuing concern for study design and implementation. This includes changes that may impact regional cropping practices, pests, yields, etc. Climate extremes (e.g., heat, cold, storms) may also have major biological impacts and make study interpretation difficult. These factors are especially important for non-*Apis* species. A risk assessment that is adequate and representative of different environments and climatic zones may be best achieved by conducting repetitive studies in different environments and climate zones.

Also the plot size should be large. The larger the plot size, the fewer issues one has to deal with when interpreting the results (i.e., determining whether the bees are visiting adjacent fields or wild areas instead of the treated crop, and if the bees are running out for forage in the treated field). One Panel member recommends a field plot size of 40 acres. In addition the design could entail having several hives used in each field. Replicate fields for each treatment are needed. However, if several hives are used per field, then fewer replicate fields would be needed than if only one hive was used per field.

The Panel recommended as a good rule of thumb that the maximum number of bees in a field trial is 30,000 adult females per acre. This recommendation is based on the alfalfa leafcutting bee in blooming alfalfa (honey bees are larger, so should not exceed this number). 30,000 bees is one moderately sized honey bee colony.

13. Risk Estimation. Sections 2 (Proposed Risk Assessment Process) and **5** (Risk Characterization) of the white paper indicate that the proposed risk assessment process is intended to be tiered and iterative. As part of the Tier I screen, a number of iterations can be conducted on exposure estimates that allow the risk quotient (RQ) values to be further refined and potentially pass the screen without requiring higher tier effects testing at the semi- or full-field level. However, while the proposed Tier I process for bees is quantitative and results in an RQ value which can in turn be compared to a Level of Concern (LOC), higher tier refinements are used to qualitatively (*i.e.,* no RQ derivation) determine whether whole hives will be adversely affected from the use of a pesticide based on the initial screening-level assessment.

a. Please comment on the use of data on individual bees to transition to higher tier studies given that the Tier I studies focus on survival as the primary measurement endpoint although additional endpoints may be forthcoming as test designs continue to develop.

Individual-level phenotype does not translate to colony-level phenotype in very effective ways. Changes in worker longevity and life-history transitions (i.e., age at first foraging) would likely be most impacted. Use of mathematical modeling would assist empirical measures greatly by generating expectations that can be verified.

Data on individual pupae can be very informative and have a greater direct impact from lower tiers to higher tiers (see Question 9 to Panel's Response). In doing so, there is a need to assess mode-of-action in order to qualitatively assess risks. For example, Insect Growth Regulators (IGRs) are not as likely to have acute toxicity, but they can also affect development or egg viability (Hodgson et al., 2011).

Higher tier studies are necessary to address risk under more realistic conditions with bee colonies and the function of hives. The function of a hive and the multiple behaviors that are needed to sustain a viable colony, the super-organism, may not be detectable in laboratory studies. Thus, if a conservative Tier I triggers higher studies because of adult or larval toxicity and behavioral aspects, higher tier tests taking into account the different important exposure routes are essential.

b. Please comment on the derivation of the Level of Concern (i.e., LOC=0.4) and the extent to which it serves as an appropriate screen to transition to higher tiers of testing/refinement.

A given LOC, however, does provide a context by which a default concentration can be tested. Essentially, EPA is proposing to bring the LOC to the NOEC. The current data is based on young worker bees. The Panel proposes including larvae and oral toxicity, so this will make the tests more sensitive.

Similarly, one basal assumption in the calculation of the LOC and its extrapolation to the field is that the compounds are used following best management practices (BMP). One of the most important values of Tier III screenings is that they may enable the capture of real-world applications including instances of non-BMP.

In the white paper, the level of concern is compared to a risk quotient that is defined as the contact or dietary exposure value divided by the contact or oral LD_{50} value. EPA proposes to set this value at 0.4 which is equivalent to approximately 10% mortality. Given the uncertainty in extrapolating effects on individual bees to effects on the colony, it is difficult to consider an appropriate level of concern for adult bee mortality at the Tier I level. As discussed in other questions, other endpoints (e.g., behavior, development, growth rate) may be more appropriate and reflective of colony health. At this point, the science does not support that 10% mortality of adult bees would result in colony collapse or even instability. Hence, the Panel believes this value is too conservative.

In order to approximate a more realistic mortality, one might look at a sensitivity analysis of mortality in currently existing population biology models such as described by Fefferman and Starks 2006.

The suggested approach seems very conservative, and is even more protective than the established Hazard Quotient (HQ) method triggering higher tier studies if the HQ is greater than 50 and seems to ensure high certainty.

c. Please comment of the quantitative aspect of the screening-level (Tier I) assessment and the use of Tier II and Tier III whole hive studies to qualitatively characterize risk.

The main utility is to provide a scale of concentrations. The field-tested concentrations, however, would need to be modulated up or down depending on effects on colony phenotype.

Different realistic exposure scenarios and different stressors are automatically included in higher tiers and are better suited to real life. As Tier I appears to be very conservative, the approach seems to ensure that higher tiers are triggered. Tier II represents a worst case which are sensitive to discover effects (confinement). Tier III represents a more realistic scenario by integrating all multiple stressors and effects that cannot be addressed by Tier II studies (e.g., homing behavior).

d. Please comment on the assumption that the effects on individual bees measured in laboratory studies must be considered in the context of whole colony studies conducted under semi-field and full-field conditions.

The Panel believes that this is the fundamental assumption that defines the problem that individual phenotype is not a proxy for colony phenotype. However because colonies are comprised of individuals, knowing both is paramount.

Depending on the nature of effects and taking into account the mode of action, the efficacy information (e.g., IGR properties), substance properties (e.g., systemic), the application method, and persistency of a compound, higher tier studies may need special adaptations to be ensure that the concerns about a substance/product are investigated.

e. Please comment on the proposed use of a weight-of-evidence approach based on information obtained from multiple tiers of risk assessment for characterizing pesticide risks to honey bees.

Using one level of biological organization as a proxy for others is always fraught with difficulty. However, some correlations are stronger than others, thus identifying those with the strongest linkages is key. In doing so, high-throughput methods that are reliable and repeatable will be most valuable. Qualitative evidence is more important in social systems than simple mortality derived from solitary systems.

The Panel also feels that the weight of evidence should be centralizing toward the Tier II level as being the most informative for social insects, because the colony is the basal unit of selection (see Question 8).

Feedback for the weight of evidence from investigations of incidents needs to be taken into account. Detecting incidents may be used to inform risk assessors and risk managers. Every single incident may not necessarily change the outcome of a risk assessment. This is because there needs to be a careful investigation of the cause (e.g., proper use/ misuse and whether there is a sound conclusion) before incidents should be taken into account for the weight of evidence. Using incident reporting may be helpful, but when considering an incident one must differentiate from the proper and improper use of a pesticide.

f. Please comment on how best to characterize overall uncertainty or weigh different areas of uncertainty in risk characterization.

Tier III studies are much more prone to a Type II statistical error (appearing to not have any significant effects, when there really are some), while Tier II studies are more prone to a Type I statistical error (appearing to have significant effects, when in reality there are none). The level of uncertainty can be determined in a Tier I type of experiment, but it is more difficult to determine in Type III studies with honey bee hives. Thus, if the test pesticide appears to have no negative effects in Tier III studies, these results need to be checked against the preponderance of evidence at lower tier levels (see also Question 13e). The number of studies with negative effects should affect the weight given to uncertainty.

g. Please comment on how to focus/prioritize uncertainties when designing and interpreting Tier II and Tier III studies.

Most of these recommendations have been addressed previously by the Panel, such as increasing sample size, increasing statistical power, including replication, factoring in genetic variability, and focusing at the Tier II level (see also Question 13d to Panel's Response).

In doing so, it is important to focus on those potential problems that are identified at lower tiers. For example, if chronic larval exposure is identified as problematic at Tier I compared to acute adult mortality, then focusing on colony-level brood measurements at Tiers II and III should take priority. Thus, focusing on a standard one or two factor/s for every study does not adequately

address the complexities of honey bee colonies; rather, those hypotheses need to be modulated according to the route of exposure and potential impact at the colony level.

One panelist notes that Bayesian methods may be useful tools for data analysis. They are designed to determine "what is the significance of what we were able to observe" rather than applying a "significant or not" test that relies on traditional, frequentist views of statistical power that may be unrealistic for studies of expensive or difficult scenarios.

14. Colony-level Modeling. As part of the proposed risk assessment process, **Section 5.4** of the white paper discusses the concept of using colony-level models as a means of integrating exposure and effects information generated from the multiple risk assessment tiers and in turn linking this information quantitatively to the proposed assessment endpoints. Conceptually, such models could inform the need for transitioning from lower tiers to higher tiers in the risk assessment process. They could also be considered in identifying critical design elements of higher tier studies (e.g., semi-filed or full field studies).

a. Please comment on the concept of using colony-level ecological models to inform the proposed risk assessment process for honey bees, as indicated above.

The Panel agrees that, in concept, employing colony-level ecological models should greatly strengthen the ability to generate valid risk assessment for honey bees. As an example, interactions and scaling questions from individual effects to colony level effects are exactly one sort of thing that mathematical models can do.

The most useful way models can accomplish this is to identify proximate, measurable outcomes that should be monitored for, rather than requiring a Boolean outcome of colony death. As long as the identified metrics are things that can actually be observed and measured, then we can use the models to identify which individual- and group-communication-level responses to pesticide exposure may lead to colony-level compromise of function without testing each type of exposure/interaction empirically.

There are a number of potential pitfalls that should be avoided. Most importantly, the Panel agrees that colony-level simulation models attempt to model, most or all aspects of colony life should not be used. As discussed/evidenced in much of the white paper discussing empirical results prior to section 5.4, critical elements on which any colony-level ecological model could be built remain unknown/uncertain. These uncertainties exist in both empirical measurements for important parameters, and in functional response relationships among variables. While the Panel agreed that models should still play a vital role in risk assessment, under these conditions, many smaller models, examining the role of uncertainty in within-colony dynamics and responses to particular parameters should be explored independently prior to inclusion in a single ecological model that is intended to capture everything. Applied mathematical models have repeatedly demonstrated how seemingly insignificant interactions among uncertain terms and functional responses can cause the predictive power of a model to drop substantially relative to the insight that can be gained by the careful consideration of the impact of uncertainties on

"smaller", more local questions, then interpreted together to inform a global result (Regan et al. , 2003; Lek, 2007).

One panel member believes that models should be used to add Tier I measures, but not to reject measures. However, more generally, the Panel agrees that, as indicated in Figure 1 of the whitepaper, risk assessment is an iterative process, and the findings from validated colony-level models can be used to evaluate the usefulness of Tier I toxicity tests and assessment endpoints in predicting colony level effects. This information can then be used to revisit, refine, and revise Tier I toxicity testing requirements, as necessary, in order to make the Tier I tests more relevant to the Agency's protection goals for bees.

The Panel generally expresses enthusiasm about the inclusion of modeling efforts as part of a thorough risk assessment. There was initial, abstract discussion about how models could be used to deal with questions of scale (i.e., how individual-level effects might interact or act in parallel to yield colony- or even population-level effects). Further, the Panel agrees with the potential for models to aid in determining when particular outcomes from spatially or temporally local studies might provide insight into broader, nation-wide or multi-year processes. (Many of these comments and discussions occurred during discussion of other charge questions, though in reference to eventual discussion during Charge Question 14 and its parts.)

The Panel expresses the view that both simulation models (as those discussed in the white paper) and analytic models would be of use to pollinator risk assessment efforts. In clarification, analytic models are those that focus on the causal relationships among variables and parameters. They use equations to capture these relationships, and are primarily used to generate insight into system-level processes (e.g., threshold behaviors for phase transitions). Examples of such phase transitions include the formation of crystals as liquids solidify, rain falling from a cloud, epidemics emerging from endemic disease spread depending on birth rates, and to end with an example from bees the convergence to consensus for a new nest site during swarming. Each of these insights into what combinations of conditions lead to the transition (i.e., liquid vs. solid, argument vs. consensus, etc.) can be made using an analytic model without the need for empirical measurement of the system. For example, there is no need for information on the number of bees arguing for each of how many different new potential nest sites to be able to use the model to determine a set of necessary and sufficient conditions for the decision to be made. Analysis can then be made about questions such as trade-offs in time vs. optimality of site quality. Similarly in the more physical examples, how to successfully seed clouds, or provide an initial lattice on which crystal formation can coalesce to produce crystals with particular properties. Basically, analytic models provide insight into how to manipulate the system, so long as the understanding of causal relationships has been appropriately captured. Peer review of analytic models focuses on appropriately capturing these causal relationships, rather than on particular values for variables or parameters, and validation refers to validation of the relationships (usually by describing conditions that imply, by logical necessity, certain behaviors, rather than comparison of predictions of particular numbers with real world measurements).

From these analytic models, one can then build simulations. In these simulations, the equations are used with particular initial variable values, and particular informed parameters to compute what are called "numerical solutions" (i.e., values predicted by the analytic equations).

Producing these values is different from analyzing the equations because it will reveal outcomes as measured values, rather than showing where those values must occur (i.e., "we tried running this model at increasing values from 0.1 to 0.9, and after we make this number bigger than 0.5, we get crystals, but not when it's less than 0.5" instead of "due to the nature of these equations, we can solve for the critical number for crystallization, and that equation will hold no matter what the parameter values for our liquid will be, so long as the relationships represented in the model are correct").

Whatever models are used, model validation involves statistical analysis of the correlation between model outcomes and real-world observations.

In choosing models to aid in pollinator risk assessment, the Panel expressed concern that, especially for Tier 1-level studies, the value of simulation models may be obviated by the need for as many measurements to parameterize them as would be needed to simply answer the question of interest. This is not the case in Tiers II and III.

Brief Examples of Analytic Models:
An example from physics can be found in the ideal gas relationship between Temperature, Volume, and Pressure expressed as PV=nRT-Where "n" is number of moles, and R is the universal gas constant. If we did not know this formula, we could experimentally determine that the relationship between temperature pressure are linear for a given volume. However then, every volume and type of gas would require validation testing to identify the constant k. Knowing the full formula provides us a way to evaluate how the relationship between temperature and volume determines pressure for every ideal gas, rather than generating a separate simulation model for each gas in order to predict the outcome. This provides insight into the existence of potential thresholds, and the values of these thresholds lead to sufficient criteria for empirical measurement to ensure that a system would remain within acceptable boundaries for application.

A honey bee example that is analogous to this system would be the use of daily worker mortality as a 'single-metric test for colony failure', as discussed during earlier Panel deliberations. However, death rates are much less informative or predictive of failure than the relationship between daily worker mortality and daily worker eclosion (the emergence of new workers). If the eclosion rate is greater than the rate that the workforce is dying, worker death is unlikely to cause failure, even if percent mortality intuitively seems relatively high. The ratio of workforce loss to workforce recruitment is what should be considered. This is an example of how an analytic model of demographic colony processes improves our understanding and reveals how the single proposed metric may fail to capture the outcome of interest.

Notice that the new metric, this ratio of birth and death rates, allows insight into whether a colony is thriving or dying without measuring a single empirical number to plug into the equations. This is the hallmark of a valuable analytic mathematical model. It provides a valuable measure that is testable. Of course, this ratio is unlikely to be all that is needed, there must also be some thresholds, such as a minimum population size for colony persistence, and potentially a maximum population size to include swarming within a colony and carrying capacity within a location.)

The Panel also agrees that analytic models can indicate "what to go measure or simulate" to understand which aspects are truly important to colony success. Models in this and other contexts have already shown cases in which intuitively important values turn out to be mostly irrelevant (Fefferman and Starks 2006), and cases in which intuitively irrelevant values turn out to be critical (Boccara 2004.).

The Panel pointed out that this does not mean there is the equivalent of a "data gap" here. One of the benefits of analytic models for this use is that all the building blocks for this type of model are already available for honey bee colony processes in various forms: growth, fission, foraging, etc. (Fefferman and Starks 2006, Khoury et al. 2011). These models can be used with the new focal questions of pesticide exposure to explore which particular elements will be most relevant to test and at which Tiers of effect.

Lastly, the Panel agreed that the Tiers themselves become problematic in a modeling context. Colony-level models can and should be used to backwards inform which individual-level measurements will be critical to colony survival. This is not a "loop" in the conceptual model sense, but changes the direction of the conceptual-flow. Thus, flow arrows. In these cases, it would be a mistake to use Tier 1 equivalent models as the first-pass studies, progressing to other Tiers only when Tier 1 results seem to indicate the need for concern. Instead, colony-level or population-level models that would mirror the types of empirical studies performed in Tiers II or III can be modeled with analytic models to determine which individual-bee tests would be likely to scale up to colony-level issues of concern. This can help directly with the relatively qualitative problem of cumulative sub-lethal effects failing to trigger Tier I progression to Tier II, while still resulting in serious compromise to colony- level success.

b. *Please comment on the state of the science regarding available honey bee models discussed in* ***Section 5.4.2*** *of the white paper in relation to their potential application in a regulatory risk assessment context. In particular, please comment on the extent that such models have been evaluated using empirical data related to honey bee population dynamics and the availability of such data for their parameterization.*

The Panel notes that many of the models discussed directly in the white paper have access to only sparse data, thus far. While the Panel agrees that it will be important to collect further input data (i.e. measurements of parameter values and valid characterizations of functional relationships), and important to collect validation data (i.e., measurements from the real world for comparison with predicted model outcomes), the Panel also noted that analytic models can be useful in the absence of such data. It would be a mistake to value the input of all models only once empirical measurements to appropriately parameterize simulation models exist. The conclusions of models can provide insight into how systems work (see discussion of analytic models above), in addition to making more specific predictions about system outcomes for focal metrics. For purposes of risk assessment, the Panel notes that this more general type of employment for models can help assess the importance of unknown threats. (e.g. In the cases in which a model demonstrates that only individual sensitivity greater than some threshold X per unit of exposure could cause the collapse of a colony by significantly altering the death-to-eclosion ratio. Even if very little is known about individual sensitivity to exposure, this can provide valuable insight.)

Further, even where data for parameterization and validation of the models have been available, these data have frequently been in aid of testing the specific predictions of each simulation model's motivating question. There is danger in assuming that a model that can consistently predict one outcome of a system is accurately capturing (and could therefore predict) all of the workings of the system without a peer review step for the analytic representation of relationship among model components. Few, if any, of these models have been compared to other models or tested for robustness of model choices apart from their specific focal questions. Model validation and calibration is also problematic since there little agreement on what constitutes an appropriate level of agreement between measured and predicted values.

Within the narrower context of models and their use for empirical simulation and prediction , the Panel notes the need to discuss appropriate sensitivity and robustness testing for individual parameters, variables, and functional relationships. These need to explicitly and separately discuss how to extract useful insight despite both biological stochasticity and measurement/relational uncertainty (this relates to Question 13.f). The Panel concluded that these discussions must be accomplished on a model-by-model basis because the impact of these different forces will vary by model design.

One panel member believes that although models may have the potential to inform risk assessment, the state of modeling is currently insufficient for practical use and that instead there should be more reliance on higher tier tests. This panel member further stresses that the current state of modeling is limited because it cannot account for all the various complexities of a bee colony.

c. *Please comment on the most important elements that should be considered in reviewing available honey bee colony ecological models for potential application in risk assessments.*

The Panel concludes that the first two questions that should be asked with any model are "What insight does it provide that accurate empirical studies could not" or, if there is none, "How does the model spare effort or expense to provide those insights." In the case of risk assessment, there are two main components: How likely is the scenario in question to occur, and how severe is the likely outcome if that scenario were to occur. For colony-level ecological models, the former question relates to exposure in individual bees and how those individual doses interact to affect single colony behavior/survival. The latter question relates to the impact of individual colony responses on the broader ecology of honey bees, and on the use of the population as effective pollinators.

The Panel further concludes that, as the models presented in the white paper were not designed with these types of questions in mind, they do not specifically address these questions in directly useful ways, but could be very important in formulating answers within each level of concern to eventually produce a useful risk assessment. However, the Panel agreed that this again points to the need for an analytic set of models to designate and integrate each of these more specific models into a broader understanding of where the focal research results for each aspect will fit into the context of broader risk assessment. Many of the panel members agreed that expanding to include this type of modeling as well will generate insight into which metrics will be most

useful, and then will allow us to generate the simulation/numerical solution models to predict actual outcomes based on values obtained experimentally.

References

Alaux, C., M. Folschweiller, C. McDonnell, D. Beslay, M. Cousin, C. Dussaubat, J. L. Brunet, and Y. Le Conte. 2011. Pathological effects of the microsporidium Nosema ceranae on honey bee queen physiology (Apis mellifera). Journal of Invertebrate Pathology 106: 380-385.

Atkins, E.L. and D. Kallum 1986. Comparative morphogenetic and toxicity studies on the effect of pesticides on honey bee brood. Journal of Apicultural Research 25: 242-255.
Aupinel, P., D. Fortini, H. Dufour, J.-N. Tasei, B. Michaud, J.-F. Odoux and M.-H. Pham-Delegue. 2005. Improvement of artificial feeding in a standard *in vitro* method for rearing *Apis mellifera* larvae. Bulletin of Insectology 58: 107-111.

Aupinel, P., D. Fortini, B. Michaud, F. Marolleau, J-N. Tasei and J-F. Odoux. 2007. Toxicity of dimethoate and fenoxycarb to honey bee brood (*Apis mellifera*), using a new *in vitro* standardized feeding method. Pest Management Science 63: 1090 – 1094.

Badano, E.I. 2011. Potential negative effects of exotic honey bees on the diversity of native pollinators and yield of highland coffee plantations. Agricultural and forest entomology 13(4): 365-372.

Boccara, N. 2004. *Modeling Complex Systems*. Springer, Chicago, Illinois

Buchmann, S. L. and G. P. Nabhan. 1996. *The Forgotten Pollinators*. Island Press/Shearwater Press, Washington, D.C.

Burken, J. G., and J. L. Schnoor. 1998. Predictive relationships for uptake of organic contaminants by hybrid poplar trees. Environmental Science and Technology. 32:3379-3385

Ciarlo, T. J., C. A. Mullin, J. L. Frazier and D. R. Schmehl. 2012. Learning impairment in honey bees caused by agricultural spray adjuvants. *PLoS ONE*, **7**, e40848-e40848.

COLOSS, 2010. Proceedings of the COLOSS Workshop "Method standardization for larval, Graz, Austria, Available online at:
http://www.coloss.org/publications/ProceedingsCOLOSSLarvalRearingGraz2010.pdf

Davis, A.R. and R.W. Shuel. 1988. Distribution of C(14)-Labeled carbofuran and dimethoate in royal jelly, queen larvae and nurse honey bees. Apicologie, 19 (1): 37-50.

Delaney, D. A., J. J. Keller, J. R. Caren and D. R. Tarpy. 2011. The physical, insemination, and reproductive quality of honey bee queens (*Apis mellifera*). Apidologie, 42, 1-13.

De Rossi, C. 2010. (Dissertation) Pesticides in the atmosphere Occurrence, distribution and behaviour of selected pesticides in the vinicultural area of Trier, Germany(http://ubt.opus.hbz nrw.de/volltexte/2010/573/pdf/DeRossiChiaraDiss.pdf).

Dohzono, I. 2010. Impacts of alien bees on native plant-pollinator relationships: A review with special emphasis on plant reproduction. Applied entomology and zoology 45(1):37-47.

EFSA. 2012. Scientific Opinion on the science behind the development of a risk assessment of Plant Protection Products on bees (*Apis mellifera, Bombus* spp. and solitary bees). EFSA Journal 2012; 10(5) 2668. [275 pp.] doi:10.2903/j.efsa.2012.2668. Available online at: www.efsa.europa.eu/efsajournal

Evans, J. D., Y. P. Chen, G. Di Prisco, J. Pettis and V. Williams. 2009. Bee cups: single-use cages for honey bee experiments. Journal of Apicultural Research, 48, 300-302.
Fefferman, N.H. and P.T.B. Starks. 2006. A modeling approach to swarming in honey bees. Insectes Sociaux 53(1):37-4

Forster et al., 2012. ICPBR-Working Group Risks posed by dusts: overview of the area and Recommendations, Julius-Kühn-Archiv, 437: 191-198

Franke. P, Juraske, R., Anton, A., Charles, R. Jolliet, O. 2012. Evaluation of Pesticides In Food-A Dynamic Multicrop Model, 6th SETAC World Congress
Available online at: http://www.lc-impact.eu/userfiles/Fantke_SETAC2012a.pdf

Frankham, R. 1995. Effective population size/adult population size ratios in wildlife: a review. Genetical Research, 66, pp 95-107 doi:10.1017/S0016672300034455.

Goulson, D. 2003. Effects of introduced bees on native ecosystems. Annual Review of Ecology and Systematics 34: 1-26.

Hendriksma HP, Hartel S, Steffan-Dewenter I 2011, Honey bee risk assessment: new approach for in vitro larvae rearing and data analysis. Methods in Ecology and Evolution 2: 509-517)

Hodgson, E. W., T. L. Pitts-Singer, and J. D. Barbour. 2011. Effects of the insect growth regulator, novaluron, on immature alfalfa leafcutting bees, Megachile rotundata. J. Insect Sci. 11: 43.

Johnson, R. M., J. D. Evans, G. E. Robinson and M. R. Berenbaum. 2009. Changes in transcript abundance relating to colony collapse disorder in honey bees (*Apis mellifera*). *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 14790-14795.

Johnson RM, Mao W, Pollock HS, Niu G, Schuler MA, et al. 2012. Ecologically Appropriate Xenobiotics Induce Cytochrome P450s in *Apis mellifera*. PLoS ONE 7(2): e31051. doi:10.1371/journal.pone.0031051

Kamel, A.; Dively, G.; Hawthorn, D. and J. Pettis. Unpublished research examining the fate of imidacloprid and its metabolites in honey bee hives.

Klein, A.-M., I. Steffan-Dewenter, and T. Tscharntke. 2003. Fruit set of highland coffee increases with the diversity of pollinating bees. Proc. R. Soc. Lond. B. 27: 955-961.

Khoury, D.S., M.R. Myerscough, and A.B. Barron. 2011. A Quantitative Model of Honey Bee Colony Population Dynamics. PLoS ONE 6(4): e18491.

Kremen, C., N. M. Williams, and R. W. Thorp. 2002. Crop pollination from native bees at risk from agricultural intensification. Proceedings of the National Academy of Sciences 99:16812-16816.

Lande, R. and G.F. Barrowclough. 1987. Effective population size, genetic variation, and their use in population management. Pages 87-123 in M.E. Soule, ed. Viable populations for conservation. Cambridge University Press, Cambridge, England.

Larney, F.J., A.J. Cessan, and M.S. Bullock. 1999. Herbicide transport on wind-eroded sediment. Journal of Environmental Quality 28:1412-1421.

Lek, S.S. 2007. Uncertainty in Ecological Models. Ecological Modelling 207, 1 1-2.

Mayer, D. F. and C. A. Johansen. 2003. The rise and decline of *Nomia melanderi* (Hymenoptera: Halictidae) as a commercial pollinator for alfalfa seed. In (K. Strickler and J. H. Cane, Eds.) For Nonnative Crops, Whence Pollinators of the Future? Thomas Say Publications in Entomology: Proceedings. Entomological Society of America. Lanham, MD. pp. 139-149.

Mullin CA, Frazier M, Frazier JL, Ashcraft S, Simonds R, et al. (2010) High Levels of Miticides and Agrochemicals in North American Apiaries: Implications for Honey Bee Health. PLoS ONE 5(3): e9754. doi:10.1371/journal.pone.0009754

OECD a.. 1998. Guidelines for the testing of chemicals Number 213 Honey bees, acute oral toxicity test, OECD, Environmental Health and Safety Division, Paris,.

OECD b. 1998. Guidelines for the testing of chemicals Number 214 Honey bees, acute contact toxicity test, OECD, Environmental Health and Safety Division, Paris,

OECD 2007. OECD Guidance Document 75. 2007. Guidance document on the honey bee (*Apis mellifera* L.) brood test under semi-field conditions. Series on testing and assessment, Number 75. ENV/ JM MONO22.

OEPP/EPPO 2001. Guideline for the efficacy evaluation of plant protection products – Side effects on honey bees.- OEPP/EPPO Bulletin, PP 1/170(3) update 2000, 19-23.

Oomen PA, De Ruijter A and Van der Steen J 1992. Method for honey bee brood feeding tests with insect growth-regulating insecticides. EPPO Bull. 22: 613-616.

Regan, H. M., H.R. Akcakaya, S. Ferson, K.V. Root, S. Carroll, L.R. Ginzburg. 2003. Treatments of Uncertainty and Variability in Ecological Risk Assessment of Single-Species Populations. Human and Ecological Risk Assessment. Vol. 9, no. 4, pp. 889-906.

Rortais, A., Arnold, G., Halm, M.P., F. Touffet-Briens. 2005. Modes of honey bees exposure to systemic insecticides: estimated amounts of contaminated pollen and nectar consumed by different categories of bees. Apidologie, 36: 71-83.

Rueppell, O., C. Bachelier, M. K. Fondrk and R. E. Page. 2007. Regulation of life history determines lifespan of worker honey bees (Apis mellifera L.). *Experimental Gerontology*, **42**, 1020-1032.

Rubik, D.W and H. Wolda. 2001. Do competing honey bees matter? Dynamics and abundance of native bees before and after honey bee invasion. Population Ecology 43(1):53-62.

Roberts, S.P. and J. F. Harrison. 1999. Mechanisms of thermal stability during flight in honey bee Apies Mellifera. J. Expt Biol. 202, 1523–1533 1999.

Schaffer, W.M. D.W. Zeh, S.L. Buchmann, S Kleinhans, M.V. Schaffer, and J. Antrim. 1983. Competition for Nectar between Introduced Honey Bees and Native North American Bees and Ants. Ecology 64(3):564-577.

Shaw, K.E., G. Davidson, S.J. Clark, B.V. Ball, D. Chandler, and K.D. Sunderland, 2002. Laboratory bioassays to assess the pathogenicity of mitosporic fungi to *Varroa destructor* (Acari: Mesostigmata), an ectoparasitic mite of the honey bee, *Apis mellifera.* Biological Control 24: 266-276.

Spannhoff, A. 1, KeeKim, Y1, Raynal, N. 2, Gharibyan, V. 2, Su, M., Zhou, Y, Li, J. 3, Castellano, S. , Sbardella, G, Issa, J& Bedford, M. 1,
2011. Histone deacetylase inhibitor activity in royal jelly might facilitate caste switching in bees, EMBO reports VOL 12 | NO 3 | 2011 &2011 European Mollecular Biology Organization

Sugden, E. A., R. W. Thorp, and S. L. Buchmann. 1996. Honey bee-native bee competition: focal point for environmental change and apicultural response in Australia. Bee World. 77: 26-44.

US EPA. 2001. GENeric Estimated Environmental Concentration Model (GENEEC), Available online at: http://www.epa.gov/oppefed1/models/water/geneec2_description.htm

US EPA 2012. EPA: Pesticides - A Discussion with the FIFRA Scientific Advisory Panel Regarding the Terrestrial and Aquatic Level II Refined Risk Assessment Models (Version 2.0) - Chapter III. Terrestrial Level II Model). Available at:
http://www.epa.gov/oppefed1/ecorisk/fifrasap/rra_chap_three.htm

Tapparo, A., D. Marton, C. Giorio, A. Zanella, L. Solda, M. Marzaro, L. Vivian, and V. Girolami, 2012. Assessment of the environmental exposure of honeybess to particulate matter containing neonicotinoid insecticides coming from treated corn seeds. Environmental Science Technology 46: 2592-2599.

Tepedino, V. J., Alston, D. G.; Bradley, B. A., Toler, T. R., and Griswold, T. L. 2007. Orchard pollination in Capitol Reef National Park, Utah, USA. Honey bees or native bees? Biodiversity and Conservation. 16: 3083-3094.

Thomson, D. 2004. Competitive interactions between the invasive European honey bee and native bumble bees. Ecology 85(2):458-470.

Thomson, D.M. 2006. Detecting the effects of introduced species: a case study of competition between *Apis* and *Bombus*. Oikos 114(3):407-418.

Tuell, J. K., J. S. Ascher and R. Isaacs. (2009). Wild Bees (Hymenoptera: Apoidea: Anthophila) of the Michigan Highbush Blueberry Agroecosystem. Annals of the Entomological Society of America, **102**, 275-287.

Wittmann D. (1982) Determination of the LC50 of Dimilin 25 WP for honey bee brood on free flying colonies as an example for the use of a new *Apis* larvae-test, Apidologie 13, 104–107

Wu, J. Y., C. M. Anelli and W. S. Sheppard. (2011). Sub-Lethal Effects of Pesticide Residues in Brood Comb on Worker Honey Bee (Apis mellifera) Development and Longevity. *PLoS ONE*, **6**.