PYRETHROID WORKING GROUP 2012

August 21, 2012

Fred Pearson (Chairperson)
Syngenta Crop Protection
410 Swing Road
Greensboro, NC 27419
Tel: (336) 632-2365
Fax: (336) 632-6374
fred.pearson@syngenta.com

Pat Devine (Treasurer)
DuPont Crop Protection
Tel: (302) 366-5274
Fax: (302) 351-7012
patricia.g.devine@usa.dupont.com

Ann Orth
FMC Corporation
Tel: (215) 299-6753
Fax: (215) 299-6456
Ann.Orth@fmc.com

Ian Chart
AMVAC Chemical
Tel: (949)221-6112
Fax: (949)476-9303
Ianc@amvac-chemical.com

Karen Shearer
Bayer CropScience
Tel: (919) 549-2365
Fax: (919) 549-3937
karen.shearer@bayercropscience.com

Diane Allemang
Pytech, c/o Cheminova
Tel: (703) 373-8883
Fax: (703) 373-8887
diane.allemang@cheminova.com

Nasser Assaf
Valent BioSciences Corporation
Phone: (847) 968-4844
Fax: (925) 817-5160
nasser.assaf@valent.com

Ms. Jeanine Townsend
Clerk to the Board
State Water Resources Control Board
1001 I Street
Sacramento, CA 95814

Public Hearing (8/21/12)
Policy for Toxicity Assessment and Control
Deadline: 8/21/12 by 12 noon

RECEIVED
8-21-12
SWRCB Clerk

Re: Comment Letter - Policy for Toxicity Assessment and Control

Dear Ms. Townsend,

The Pyrethroid Working Group (PWG) is an industry-based group that works together on regulatory issues associated with uses of pyrethroids as pest-control agents used in crop protection, residential settings and public health. The group comprises eight companies who are basic manufacturers and primary registrants of a number of pyrethroid insecticides.

The PWG appreciates the opportunity to comment on the State Water Resources Draft Policy on Toxicity Assessment and Control. We offer these comments as a supplement to the previous review comments prepared on a draft of this document by the Western Plant Health Association (WPHA) dated January 21, 2011, and have included Amendment A from the WPHA comments. The previous policy document reviewed in 2011 is very similar to the current 2012 document so it appears that the comments previously submitted by WPHA were not taken into consideration.

We are concerned that the statement of the null hypothesis overturns years of precedent in toxicity testing by requiring the discharger to demonstrate that the effluent is not toxic rather than requiring a demonstration that the effluent is toxic. The Test of Significant Toxicity is intended to overcome common criticisms of ANOVA to indicate effects at a given concentration, but it is very likely that the policy will cause additional testing solely as a result of variability in test results rather than as a result of toxicity to the test species. The TST methodology requires debatable assumptions about power to detect effects, magnitude of effect, and error rates. While the TST method attempts to correct the deficiencies of the current methods, using a 25% effect (<75% of control) as a regulatory threshold assumes characteristics of the test results. For example, the magnitude of acceptable effect should be higher for plants than for animals.

Our analysis shows quite clearly that the reversed null hypothesis results in a level of protection far beyond a 20 or 25% effect as intended by the definition of the Regulatory Management Decision. The difficulties are illustrated in several ways in the following discussion.

Syngenta Crop Protection
410 Swing Road
Greensboro, NC 27419

DuPont Crop Protection
Stine-Haskell Research Center
PO Box 30
1090 Elkton Road
Newark, DE 19714

AMVAC Chemical
4695 MacArthur Ct.
Suite 1250
Newport Beach
California, 92660

Bayer CropScience
2 T. W. Alexander Dr.
P.O. Box 12014
Research Triangle Park
North Carolina, 27709

FMC Corporation
Agricultural Products Group
1735 Market Street
Philadelphia, PA 19103

Pytech, c/o Cheminova
Cheminova, Inc.
1600 Wilson Boulevard
Suite 700
Arlington, VA 22209

Comment #1

From the State Water Resources Board *Policy for Toxicity Assessment and Control* we have this definition:

Regulatory management decision (RMD) is the decision that represents the maximum allowable error rates and thresholds for chronic and acute toxicity (and non-toxicity) that would result in an acceptable risk to aquatic life.

The decision criterion assumes that the effluent is toxic and specifies that effects as large as the RMD should be permitted.

Chronic Toxicity
The chronic toxicity objective is expressed as a null hypothesis and a regulatory management decision (RMD) of 0.75 for chronic toxicity methods, where the following null hypothesis shall be used:

Ho: Mean response (IWC) < 0.75 • mean response (control)

Attainment of the water quality objective is demonstrated by rejecting this null hypothesis in accordance with the statistical approach described in Appendix A.

Acute Toxicity
The acute toxicity objective is expressed as a null hypothesis and an RMD of 0.80 for acute toxicity methods, where the following null hypothesis shall be used:

Ho: Mean response (IWC) < 0.80 • mean response (control)

Attainment of the water quality objective is demonstrated by rejecting this null hypothesis in accordance with the statistical approach described in Appendix A.

This stipulates the Chronic effects up to 25% and Acute effects up to 20% are permitted.

The details for the Test of Significant Toxicity (TST) approach described in the above document are discussed in detail in paper by Denton et al. (2011) recently published in Environmental Toxicology and Chemistry (Volume 30: 1117 – 1126). Therefore, the Denton et al 2011 paper was reviewed.

Figure 3 from Denton et al (2011) shows that if an applicant has an effluent that is performing at the RMD level of producing a 25% effect (vertical green dashed line in figure), then the probability of the effluent being declared toxic is between 0.8 and 0.95 depending on the alpha-level of the statistical test. This shows that in reality a 25% effect is NOT allowed under this procedure. Even a 20% effect has a probability greater that 0.4 of being declared toxic. To have less than 5% risk of having the effluent being declared toxic, then the effect must be less that 10% if the CV= 0.1, less than 5% if the CV = 0.15, essentially 0% if the CV = 0.2, and if the CV = 0.25, then the effluent will have to perform BETTER than the control.

These results show that this procedure is not allowing a 25% RMD as stated in the policy document. The degree to which the effective RMD is less than 25% depends on the variability of the data. At high levels of variability this procedure may require the effluent to perform better than the control to be assessed as nontoxic.

Comment #2

Figure 3 from Denton et.al. (2011) makes clear that the effective RMD is not the same as the coefficient of the mean response (control) as stated in the policy document. However, Figure 3 does show that by adjusting this coefficient it is possible to achieve an acceptable risk of being declared toxic at the RMD. For example, for a CV = 0.1, if the coefficient of mean response (control) were changed from 0.75 to 0.6, then the probability of being declared toxic if the applicant effluent had a 25% effect would be less than 0.05. If the CV is greater than 0.1, then the adjustment will have to be larger. The most reasonable way to accomplish this is to fix the coefficient at a level that results in a reasonable effect RMD when variability is low but achievable.

Specific comments

1. Page 6, paragraph 4 – The salinity threshold (1 ppt or greater) used for testing marine organisms is very low. Marine organisms, typically found at 35 ppt, may be stressed by oligohaline conditions. It is not clear how controls will be handled in this situation. For example, if the ambient water tested is 2 ppt salinity and the marine test species are held at 35 ppt, an salinity acclimation procedure should be used.
2. Page 15, Appendix A, Step 1 – Step 1 states that "Prior to analysis, if the measured response is reported as a percentage (e.g., percent survival, percent fertilization) it must be transformed using the arc sine square root transformation". The transformation described is appropriate for percentage data derived from a binomial experiment where n organisms or other experiment units are used in a test and binary responses (e.g., live/dead) are obtained for each experimental unit. It is not appropriate for percentage data derived from continuous measures such as percent reduction in weight and is not appropriate for non-percentage data. Specifying that the arcsine square root transformation methodology must be used limits the opportunity to use more modern methods for analyzing binary response data such as General Linear Models with random effects. It even limits the use of other traditional approaches such as logit or probit transformations which might be more appropriate

Amendment A

The power to detect an effect of a given size should be specified as well as specific statistical tests. The size of the sample on which it is based, the variability of the response across samples, and the statistical significance of the comparison are all critically important. The proposed method places a premium on one calculation without regard to sample size, sample-to-sample variability, power, or statistical significance.

The additional references cited provide some of the missing detail but do not alter the fact that a single observation is used to determine whether additional testing is required. This deficiency is partly addressed in section 6 Statistical Method, page 7, where Welsch's t-test is indicated, and Table 1, page 9, where different requirements are given on a per-species basis.

There is difficulty with the description of the test given in steps 4 and 5, page 8. The test statistic is given as

$$t = \frac{\bar{Y}_t - b \bullet \bar{Y}_c}{\sqrt{\frac{S_t^2}{n_t} + \frac{b^2 S_c^2}{n_c}}}$$

Where $\bar{Y}_c$ and $\bar{Y}_t$ are the mean responses in the control and IWC, respectively, and b is the specified proportion (0.75 for chronic tests and 0.8 for acute tests). This test statistic, t, will be positive if the response at the IWC exceeds 75% of the control mean, which means the sample passes according to the earlier cited text. However, the instructions are to compare the value t to the appropriate positive critical value in Table 2 and fail the test if t does not exceed the critical value. This test assumes the contaminant is toxic unless the data demonstrate otherwise, the opposite of most toxicity tests, where a compound is assumed non-toxic unless the data demonstrate otherwise. Depending on variability, this may not be reasonable for the applicant, as sample sizes to achieve sufficient power to demonstrate non-toxicity can be large. Two examples are provided below for Daphnia magna reproduction and Selenastrum growth where under typical conditions this approach will have adequate power under normal experimental design conditions but not under high variance scenarios that may be encountered in routine testing. Data are not available to determine whether the criteria in Table 1 are reasonable for most other species.

There is an error in the methodology for handling percent effects, such as survival. Step 1 on page 7 indicates that percent effect should be changed to proportion effect and then transformed by the arc-sine square-root transform before calculating the test statistic t shown above. While there is no disputing the value of such a transform, the formula for t is not correct for transformed data. This is because a 100b% effect in the response (percent or proportion survival) is not the same as a 100b% effect in the transformed response. Nor is there a simple fix. If the percent survival in the control is 100% and the number of observations per rep is 10, then the value of $\bar{Y}_c$ is 1.5708 - arcsin(square-root(1/40))=1.5458. If the desired effect to find is 80%, such as for fathead survival, then 0.8* $\bar{Y}_c$ =1.237. But a back transform of 1.237 is 0.89, so the test compares the observed treatment effect to an 11% mortality rate, not a 20% rate, a much more severe restriction. If the number of observations per rep is 20 instead of 10, then the maximum passing observed treatment mortality rate is 15%. If the number of observations per rep is 5 then only a 5% or lower observed mortality passes.

Furthermore, if the control survival is 90%, then a 20% reduction in the control survival is a survival rate of 72%.. However, a 20% reduction in arc-sin square-root of 0.9 is 1.237 which back-transforms to 0.89, only 1% more mortality than in the control, so the formula is much more restrictive than the nominal value and it becomes almost impossible to pass. These calculations are refined below in the example for fathead survival.

The problem indicated for survival responses is a simple example of the broader problem of computing a p% effects concentration using a transformed response. There is no simple solution.

Power Calculation Example:
Daphnia Reproduction (TYS21). For routine studies, Vrep ranges from 100 to 238 and the control mean ranges from 74 to 161. From Table 1, the appropriate chronic effect is 25% and the false negative rate is 20%. The critical value is dependent on the number of replicates, or more specifically, on the degrees of freedom of the t-test. Assuming homogeneous variances and common number, r, of reps in treatment and control, the formula for t is

$$t = \frac{\bar{Y}_t - 0.75 * \bar{Y}_c}{\sqrt{1.75 * \frac{Vrep}{r}}}$$

The following table was constructed using the mean value, 117.5, for the control mean, and three values, 100, 169, and 238 for the variance, representing the minimum, mean, and maximum observed variances.

Table 3: Daphnia Magna Reproduction: Maximum Observed Effect to Pass

vrep	reps	maxeff		vrep	reps	maxeff		vrep	reps	maxeff
100	3	9		169	3	5		238	3	1
100	4	13		169	4	10		238	4	7
100	5	15		169	5	12		238	5	10
100	6	16		169	6	14		238	6	12
100	7	17		169	7	15		238	7	13
100	8	17		169	8	15		238	8	14
100	9	18		169	9	16		238	9	14
100	10	18		169	10	16		238	10	15

Thus, under the minimum variance scenario (VREP=100), the observed effect at IWC cannot exceed 9% of the control mean to pass if there are 3 reps per treatment and control. With 10 reps in each group, the observed effect at IWC cannot exceed 18%. Under the maximum variance scenario, if there are only 3 reps in each group, then any effect exceeding 1% at IWC will fail. Again under the maximum variance scenario, if there are 10 reps per group, then any observed effect exceeding 15% will fail.

Table 4: Selenastrum Growth: Maximum Observed Effect to Pass

cmean	vrep	reps	maxeff		cmean	vrep	reps	maxeff		cmean	vrep	reps	maxeff
1.2	0.0007	2	11		1.2	0.003	3	16		1.2	0.007	3	12
1.2	0.0007	3	21		1.2	0.003	4	18		1.2	0.007	4	15
1.2	0.0007	4	22		1.2	0.003	5	19		1.2	0.007	5	17
1.2	0.0007	5	22		1.2	0.003	6	20		1.2	0.007	6	18
1.2	0.0007	6	22		1.2	0.003	7	20		1.2	0.007	7	18
1.2	0.0007	7	23		1.2	0.003	8	21		1.2	0.007	8	19
1.2	0.0007	8	23		1.2	0.003	9	21		1.2	0.007	9	19
1.2	0.0007	9	23		1.2	0.003	10	21		1.2	0.007	10	19

For Selenastrum, 3 reps per group is typical, so that a maximum observed effect at the IWC that will pass is 12% under the high variance scenario and 21% under the minimum variance scenario. The variances and means are typical of routine testing.

Fathead Survival

Fathead survival was examined assuming 10% control mortality and otherwise following the guidelines of Table 2. Table 5 summarizes the power properties. The table shows only combinations of number of fish per group, number of reps per group, and maximum percent reduction from control mean will pass the criteria. Since there is no entry with 4 reps and sample size 10, it should be inferred that for such a design, the test will invariably fail. With 5 reps of size 10, the test will fail if the observed mortality in the IWC group exceeds that of the control by more than 4%. Since 4% of 40 is 1.6, this means that if more than one additional fish dies in the treatment group beyond what die in the control, the test will fail. This is a severe failure criterion. With 4 reps of 20 fish each, the maximum increase in mortality over the control is 7%. Since 7% of 40 is 2.8, this means if 3 or more fish die in the treatment group over the number of control mortalities, the test will fail. The stated failure criteria for percent effects appear too strict to be of practical importance and will trigger further testing routinely.

Table 5: Fathead Survival: Maximum Observed Effect to Pass

p0	pt	n	b	reps	maxeff		p0	pt	n	b	reps	maxeff
0.1	0.2	5	0.8	8	2		0.1	0.2	15	0.8	7	9
0.1	0.2	5	0.8	9	3		0.1	0.2	15	0.8	8	10
0.1	0.2	5	0.8	10	4		0.1	0.2	15	0.8	9	11
0.1	0.2	10	0.8	5	4		0.1	0.2	15	0.8	10	11
0.1	0.2	10	0.8	6	5		0.1	0.2	20	0.8	3	1
0.1	0.2	10	0.8	7	7		0.1	0.2	20	0.8	4	7
0.1	0.2	10	0.8	8	8		0.1	0.2	20	0.8	5	9
0.1	0.2	10	0.8	9	8		0.1	0.2	20	0.8	6	10
0.1	0.2	10	0.8	10	9		0.1	0.2	20	0.8	7	11
0.1	0.2	15	0.8	4	5		0.1	0.2	20	0.8	8	12
0.1	0.2	15	0.8	5	7		0.1	0.2	20	0.8	9	12
0.1	0.2	15	0.8	6	8		0.1	0.2	20	0.8	10	13

Sincerely,

Fred Pearson,
Chairperson, Pyrethroid Working Group