



Department of Pesticide Regulation



Paul Helliker
Director

MEMORANDUM

Gray Davis
Governor
Winston H. Hickox
Secretary, California
Environmental
Protection Agency

TO: Joseph Frank, Senior Toxicologist
Worker Health and Safety Branch **HSM-03008**

FROM: Sally Powell, Senior Environmental Research Scientist
Sheryl Beauvais, Staff Toxicologist [original signed by S. Powell & S. Beauvais]
Worker Health and Safety Branch

DATE: February 13, 2003

SUBJECT: REVIEW OF FINAL DRAFT ARTF REPORT ON USE OF TRANSFER
COEFFICIENT AS A GENERIC TOOL

We have reviewed the final draft of the Agricultural Reentry Task Force (ARTF) report “Transfer Coefficient: A Generic Tool for Estimating Pesticide Exposure to Agricultural Workers Who Re-Enter Treated Crops” (Bruce *et al.*, 2002).

Major criticism:

The stated purpose of the investigation is to assess the generic nature of the transfer coefficient (TC). The report examines several dozen dislodgeable foliar residue (DFR) and exposure monitoring studies for evidence of the generic nature of the TC, and while it discusses the effects of numerous factors on various outcomes, it does not systematically address the stated purpose.

The generic nature of the TC has typically been assumed. Fundamental assumptions underlying the TC concept are that reentry exposure is proportional to DFR (i.e., is a linear function with zero intercept), and that for any crop-activity pair, the proportionality constant (i.e., the slope of the linear function or the TC) is the same regardless of DFR level, active ingredient (AI) or most other conditions.

In order to assess the generic nature of the TC, the fundamental “generic-ness” assumptions must be tested. The questions that need to be asked are:

1. What is the function relating exposure to DFR for each AI/crop/activity combination?
2. Does the same function describe the relationship for every AI/crop/activity? for any AI/crop/activity combinations that are expected to form a cluster?
3. Does the function fit equally well for each combination in a proposed cluster?
4. Do other factors (for example, specific crop characteristics, specific AI characteristics, age of residues, measurement methods, foliar moisture) affect the relationship between exposure and DFR within clusters? And if so, is the effect large enough to make using a generic TC undesirable?



The ARTF report does not address the first three questions at all. It simply assumes that the function relating exposure to DFR is always linear with zero intercept. Goodness-of-fit of the function is not discussed. The report attempts to address the fourth question by discussing the effects of some of these factors on within-group CVs for DFR and/or exposure, but directly examines effects on the TC for only a few.

The report fails to fulfill its purpose because most of the reported analyses miss the point. They provide some interesting and valuable information, but they do not address the stated purpose of the investigation. Specific analyses are discussed in the comments that follow.

Specific comments:

Section 2. Methods (p. 15)

Comparisons summarized in the report were performed using data from an Access database compiled from studies purchased and conducted by ARTF (p. 15). TCs were calculated according to ARTF SOPs; the procedures were not further described in the report. Because decisions made about handling data and calculating TCs can affect the results, the SOPs should be provided as an appendix.

The definition of an exposure dataset as all “measurements for one body part from all the workers monitored on one day of a study” (p.15) is perplexing. Exposure of the whole person should be the focus. Exposure to specific body parts (aside from head/neck and hands) is usually not of practical interest, even in chemical-specific exposure monitoring studies. Body part-specific TCs (except, sometimes, for hands) are not used. An exposure dataset should therefore be all measurements of total-body (minus head and hand) exposure from all the workers monitored on one day of a study.

It requires a great deal of effort to understand the assembled data using the dataset summaries in Appendices A and B. Moreover, the appendices do not indicate whether the same or different workers were monitored on different days of a study. A presentation like Table A (attached) would be very useful for understanding the design, the actual numbers of replicates and other characteristics of each study important to the analysis of generic-ness. It could readily be seen why particular studies are used for particular analyses and what the sources of variance are in those studies. This or a separate table should summarize how many datasets of each type were omitted from the comparisons because they met the exclusion criterion of having most of their values be below the LOQ (p. 16).

Section 3. Statistical Consideration (p. 17)

The text appropriately discusses the random effects model and the estimation of variance components. However, with a few exceptions, the analyses reported throughout do not apply this statistical model.

Section 4. Variability in Measurements of Exposure and DFRs (pp. 18-34)

The purpose of the information presented in subsections 4.1- 4.8 (Tables 2-11) is not clearly explained and the presentation may be misleading. Headings such as “Variability Among Monitoring Methods” (subsection 4.1) imply that differences among, e.g., DFR measurements made by different methods are being presented. “Variability Between ARTF and Purchased and Member Studies” (Table 3) might seem to present variability between results of the three types of studies. However, none of the tables in Section 4 shows variability *between* conditions. They show the average *within-dataset* variability under each condition. While comparable within-dataset variability is a prerequisite, it does not establish generic-ness, because it does not rule out, e.g., that different measurement methods yield significantly different mean DFR levels. The more important question is whether factors like AI, measurement method, etc., contribute significantly to the variance of TC. If they do, doubt is cast on the generic nature of the TC.

No unified statistical approach is applied to estimating and testing differences between variances under different conditions. It is not very meaningful to compare within-group variances one factor at a time. To see why, consider the comparison of within-dataset variability of exposure measurements obtained by whole-body and patch dosimetry (Table 2). If the whole-body and patch dosimetry studies in the database tended to use different AIs, crops, leaf textures or activities, and any of those factors contributed to differences in exposure variability, any apparent difference between methods could actually be due to the other factor(s). And, apparent similarities between methods could be due to masking by the effects of other factors. Even the comparison of inner to outer dosimeters may be confounded in this way, because the studies do not overlap completely.

The rationale for the DFR categories used in Table 9 is not given. Since the grouping can affect the results, it should be justified. Further, since an effect of DFR level on exposure variability may occur only at the extremes, it would be preferable to subdivide the lowest and highest DFR categories. A better alternative, which would avoid the problem of grouping, would be to simply plot the coefficient of variation (CV) for each study against DFR level.

A definition of “aged” residues (p. 26) should be provided. Generally, when one is talking about “aged” residues, the implication is that residues are covalently bound to a matrix (e.g., sediment, cholinesterase), and therefore are less available than before. Does ARTF possess data suggesting that “aged” residues are less available to the worker (or to routine analytical extraction)?

The odd definition of exposure “dataset” used in this report further undermines the value of these tables. It is not apparent why the average within-set variability in a group including sets of leg patches, sets of arm patches and sets of other body-part patches is meaningful. Again, the total body exposure is of interest, not the individual parts.

Section 5.1. Dislodging Techniques (pp. 35-37)

This section reports an analysis comparing measured DFR with four different dislodging techniques, for two chemicals and two crops. We wonder how meaningful this comparison can be, since it used DFR samples collected 1 day after actual applications. It seems that observed differences could be due to variation in deposition and early dissipation rates, as well as to dislodging technique. The analysis-of-variance table, means comparisons tests and the cell means and sample sizes should all be reported.

Data from the ARTF study of variation in DFR by dislodging techniques were not presented in this report (pp. 36 – 37); possibly, they were published elsewhere (e.g., Bruce, 2001). If so, that reference should be cited.

Treatment 4 (dislodging once with NEKAL solution) yielded DFR results that were significantly lower, and treatment 2 (dislodging x 3 with OT-75) yielded results that were significantly higher than the DFR technique used in most studies in the database (p. 37). No information is given about whether these data were excluded, based on the significant difference. If not, then purpose of this comparison is hard to understand.

Section 5.2. Plot-to-Plot variation in DFRs (pp. 37-41)

This section is mystifying. It talks about variation between plots within the same field, but nowhere in the description of the study or the data (Table 13) do we find anything indicating there were multiple plots per field. This needs to be clarified. This section is one of the few cases where a statistical model is actually applied, and also where more than just within-dataset variability is examined. It would be good to give the analysis of variance table, for this and each other analysis whose results are discussed. (They could be in an appendix.)

In addition, an atypical study was used for this comparison, in that half of each DFR sample was collected from a section of the field that had not been irrigated, and the other half was collected from an irrigated portion of the field (p. 37). Although it's noted that the effect of compositing disparate samples in this manner is to decrease plot-to-plot variation (p. 40), these data are used to reach a general conclusion about the relatively small effect of between-field (or plot-plot) variation (p. 41). Such a generalized conclusion can't come from such anomalous data.

Section 6, Generic Nature of Initial DFRs (pp. 41-47), and Section 7, Leaf Texture: Effect on Initial DFRs for Specific Chemicals (pp. 47-74)

These two sections represent a diversion from the stated purpose of the report. It has always been intended that generic TCs, if they proved to be so, would be used in conjunction with chemical- and crop-specific DFRs to predict future exposures. In these sections, ARTF proposes to test the hypothesis that initial DFR, normalized for application rate, is independent of AI, application method, crop, leaf type, etc. If this is the case, "then a generic value could be used as

a reasonable worst-case value for DFR in exposure assessments, or to determine whether additional studies are warranted for a given active ingredient” (p. 41).

The fits of the regression models relating initial DFR (IDFR) to application rate (AR) are not spectacular. The overall R^2 was 0.47; in 22 subsets of the observations, R^2 ranged (omitting the single lowest and highest R^2 values) from 0.02 to 0.62, and the predicted IDFR (again omitting the lowest and highest) ranged from 0.62 to $1.73 \mu\text{g cm}^{-2}/\text{lb ac}^{-1}$. That much difference in DFR could translate to a considerable difference in exposure. If generic IDFRs were to be used, the Department of Pesticide Regulation would probably have to use an upper-bound value.

One reason the regression fits are not better may be the way IDFR is quantified. Measured IDFRs tend to be unstable. DFR typically drops rapidly right after application, so differences of a very few hours in the time post-application of IDFR sampling could greatly affect the measurements. In addition, very early IDFR sampling may encounter nonuniformly dried residues, which may contribute to variability. Finally, for reasons that are unclear, it is not uncommon to see an increase in measured DFR after the initial day. Because of this instability, the Worker Health and Safety Branch does not use measured IDFRs. Instead, IDFR is estimated as the intercept of the dissipation curve (Edmiston *et al.*, 2002; also see the model on p. 39 of the ARTF report).

Some specific questions about this section follow:

1. Why was AR log-transformed in the regression model? Is there any evidence that AR is lognormally distributed?
2. A reference is needed for the statement that airblast is a less efficient method (i.e., a smaller proportion of applied pesticide reaches foliage) than ground boom or aerial applications (p. 49).
3. On pp. 52 – 53, height of tomato plants is used as a surrogate measure for foliar surface area. It's reasonable to expect a taller crop to have more and larger leaves, but perhaps not to expect linear relationships between crop height and leaf surface area as implied in the Chemical A DFR comparison among applications to tomatoes.
4. DFR data provided for Chemical B included three studies in which the study location (state) was not specified (p. 54). That information should be provided for all studies, as climate greatly impacts DFR dissipation (and apparently initial DFR, based on discussion on p. 52).
5. Figure 8 shows initial DFR mean and range for Chemical F in two crops (p. 60). How many applications were made to each crop?
6. In Table 25, a definition should be given for “% Foliage” (p. 67).

Section 8. Chemical Type: Effect on TCs (pp. 74-78)

This is the first section that directly addresses the purpose of the investigation. It compares TCs for two chemicals measured in the same DFR samples. The two chemicals are a parent

compound and degradate. It would be helpful to have reported the days post-application of the DFR and exposure monitoring, since this may affect the parent-degradate relationship.

Furthermore, how often were exposure measurements for degradates below LOQ? Were LOQs identical for parent and degradate in each case?

It makes sense that highly water-soluble compounds would yield higher DFR estimates than insoluble compounds, as mentioned on p. 77 (DFR is dislodged with an aqueous solution). This suggests that studies designed to estimate TCs should always use compounds with low water solubility. Otherwise, exposures to insoluble compounds may be underestimated if TC are derived from highly soluble compounds.

These results are interesting, but it's unclear how they would generalize to unrelated chemicals. It would be very useful to have studies comparing TCs for two chemicals applied in the same tank mix.

Section 9. Timing of Monitoring Relative to Application: Effect on TCs (pp. 78-88)

Subsection 9.1 examines the within-study variability of TC by day after application (DAA). As discussed with respect to Section 4, equal within-group variability is important, but is not the outcome of primary concern.

Subsection 9.2 examines the more critical question, the magnitude of TCs across reentry days. However, the analysis is not adequate to show there is no effect of DAA. First, the studies do not look at changes in TC over sufficiently long intervals. Most of the studies measured TC on three consecutive days, a span over which there might not be any change. Only eight studies looked at nonconsecutive days; in three of these, the comparison is confounded by moisture differences (which are discussed in Section 10), and two others have no early monitoring (AR1014 begins on Day 34 after application, AR1015 on Day 14), leaving only three studies on which to base the comparison. While there is no apparent effect of DAA on TC in these studies, this is very little data to support a conclusion. Another study, ARF012, would have provided useful data for this comparison (exposures were monitored on Days 4, 6 and 33), if the Day 33 data had not been excluded. Data from Day 33 were omitted simply because they yielded a very large TC (Klonne *et al.*, 1999).

Furthermore, looking only for a linear change in TC is too narrow. The relevant question is whether there are any differences between TCs obtained soon after application and those obtained later. A better approach would be to do analysis of variance on the individual TC values by day. This would detect any differences, not just linear trends, and *post hoc* comparisons could be used to identify the specific differences. (This analysis also permits testing the equality of variance of TC across days.) Even if there were no increasing or decreasing trend, a significant effect of DAA would signal a problem with the generic TC concept.

The analysis reported in subsection 9.3 is almost the same as that proposed in the preceding paragraph. The difference is that what we propose presumes DAA is a fixed effect, while the ARTF authors treat days as random. Although DAA actually is fixed by experimental design, researchers typically treat it as random. And since in the ARTF report, DAA is consistently treated as a random effect, it should be here as well. Unfortunately, that gives no way to test for differences between specific days. The analyses in 9.3 find that DAA contributed significantly to TC variability in 18 of 31 studies. These individual findings have to be interpreted with attention to the actual DAAs monitored (per the discussion in the preceding paragraph), the dissipation rate of the AI, and whether there were moisture present any of the monitored days.

We do not understand the final conclusion in subsection 9.3: “day-to-day effects seem to have a larger impact on the variation of TDE TCs than worker-to-worker effects. This supports the generic nature of TCs relative to timing of reentry after application” (p. 88). We would conclude the opposite.

Section 10. Moisture During Exposure Monitoring: Effect on TCs (pp. 89-95)

As presented, the results on the effects of moisture suggest an increase in TC due to moisture during monitoring, but are not easily interpreted nor integrated. There is potential confounding with DAA or DFR level. One suggestion for integrating the results is to 1) categorize moisture levels as Dry, Light and Heavy, and 2) express TC for each moisture level as a percent deviation from its study mean TC. Figures A and B in the attachments show mean percent deviation by day after application. The effect of moisture might be tested using a general linear model with day after application as a covariate. Sally did some analyses of this type, which suggested that there is a moisture effect on Dermal as well as Potential TC, but that it may only be at the heavy moisture level.

The effect of moisture on DFR could reasonably be expected to be related to water solubility of the AI. The AIs considered on pp. 89 – 95 include a relatively narrow range of water solubilities. All were in the ranges characterized in this report as insoluble to moderately soluble: chlorothalonil (0.81 mg/L @ 25°C); cyfluthrin (2 mg/L @ 20°C); ethyl parathion (24 mg/L @ 25°C); carbaryl (120 mg/L @ 20°C); malathion (145 mg/L @ 25°C). None of the AIs were highly soluble. This limitation should be mentioned as part of the conclusion.

Appendix C: The Statistical Framework for Transfer Coefficients (p. 191)

It seems that Eq. (7) should distinguish between the “true” cluster, a hypothetical construct, and the cluster that consists of all the fields that were measured. If $\log(\mu_s)$ were the mean of all fields in the hypothetical cluster, then Eq. (7) might be

$$\log(\Delta_f) = \log(\mu_s) + (\log \mu - \log \mu_s) + (\log \Delta_f - \log \mu).$$

Otherwise, the fields would be fixed effects, which doesn't seem like what you want.

Other comments:

Some references cited in the document were not provided in the reference list:

- Fuller *et al.* (2001) – p. 11
- Adams (1984) – p. 11
- Rosenheck *et al.* (2001) – p. 21
- EPA (1984) – p. 35 (also cite (p. 35) and add reference for EPA, 1997)

A list of abbreviations would be helpful to the reader, as abbreviations are defined only at first use and this is a long document.

Recommendations:

It seems to us that the ARTF database project has two phases that need to be conducted separately. The generic nature of the TC must be established, and then a generic TC database can be constructed. Instead, it seems that ARTF has constructed a database and is now trying to use it to establish generic-ness. The fact that the studies in the database do not lend themselves well to this purpose is reflected in the report by the obvious difficulty of meaningfully testing the effect of any one factor unconfounded by a host of others.

Some of the important questions about generic-ness may only be answerable with controlled experiments. For example, chemical specificity of TCs would best be addressed by applying multiple chemicals in the same tank mix. If the same mix were applied to different crops at the same time in the same region, crop-by-chemical interactions could be studied unconfounded by weather conditions. Applying the same AI to adjacent plots at different rates could disentangle the effects of DFR level and residue age. Some questions may be possible to answer adequately using the database, but the analyses must better address the confounding of effects in these studies.

This investigation gave too much attention to secondary questions and neglected the major issue, the robustness of the exposure-DFR relationship to AI, age of residue, moisture, crop and leaf type, etc. Modeling the exposure-DFR relationship, in pooled studies, using general linear models with terms for these factors and their interactions with DFR level would be a unified approach to finding the best function for the exposure-DFR relationship, testing the generic nature of the TC, and identifying clusters with common exposure-DFR functions.

References:

- Bruce, E.D. 2001. Evaluation of Dislodgeable Foliar Techniques, Statistical Distributions, and Initial Residues. ARTF Study Number ARF054. Report dated December 18, 2001. Sponsored by Agricultural Reentry Task Force, LLC.

Joseph Frank
February 13, 2003
Page 9

Bruce, E.D., Holden, L.R., and Korpalski, S. 2002. Transfer Coefficients: A Generic Tool for Estimating Pesticide Exposure to Agricultural Workers Who Re-Enter Treated Crops. Final Draft Report dated October 18, 2002. Sponsored by Agricultural Reentry Task Force, LLC.

Edmiston, S., Powell, S., Spencer, J., and Curtis, C. 2002. Guidance for Determination of Dislodgeable Foliar Residue. HS-1600. Sacramento, CA: Worker Health and Safety Branch, Department of Pesticide Regulation, California Environmental Protection Agency.

Klonne, D.R., Artz, S.C., Prochaska, C., and Rotondaro, A. 1999. Determination of Dermal and Inhalation Exposure to Reentry Workers During Harvesting in Cauliflower. ARTF Study Number ARF012. Report dated September 8, 1999. Sponsored by Agricultural Reentry Task Force, LLC.

Attachments

cc: Christine Norman, Health Canada PMRA
Dave Johnson, ARTF
Jeff Dawson, U.S.EPA OPP
Jeff Evans, U.S.EPA OPP
Mary Mitchell, Health Canada PMRA

Table A. Designs of studies in the ARTF database (cell entries are numbers of samples for DFR, number of workers for exposure).

		DAA											
		0	1	2	3	5	7	8	10	14	21		
AR1002	DFR												
	Peach orchard crop smooth leaf	Plot a (+ Days 28,35,42,49,56,63)	4-hr 2	2	2	2		2			2	2	
		Plot b	4-hr 2	2	2	2		2			2	2	
	Foliage moisture (NA)												
Vinclozolin EC 0.9 lbs/ac mod sol	Exposure (WBD)												
	AM*	5 ^a								3 ^a			
	PM*	5 ^a								5 ^a			
	* same individuals												
AR1003	DFR												
	Apple orchard crop smooth leaf	Plot a	3-hr 3 8-hr 3 12-hr 3	3	3	3	3	3			3		
	Triadimefon WP 1.0 lbs/ac mod sol	Plot b	3-hr 3 8-hr 3 12-hr 3	3	3	3	3	3					
		Plot c (+ Days 28,35)	3-hr 3 8-hr 3 12-hr 3	3	3	3	3	3			3	3	
		Foliage moisture (NA)											
	Exposure (WBD)		10 ^b							10 ^b			

a Same individuals monitored on these days of the exposure study.

b Partial overlap of individuals monitored on these days of the exposure study.

Note – information italicized in this table was invented for the purpose of completing the table.

Fig. A Change in Potential TC by moisture level during monitoring

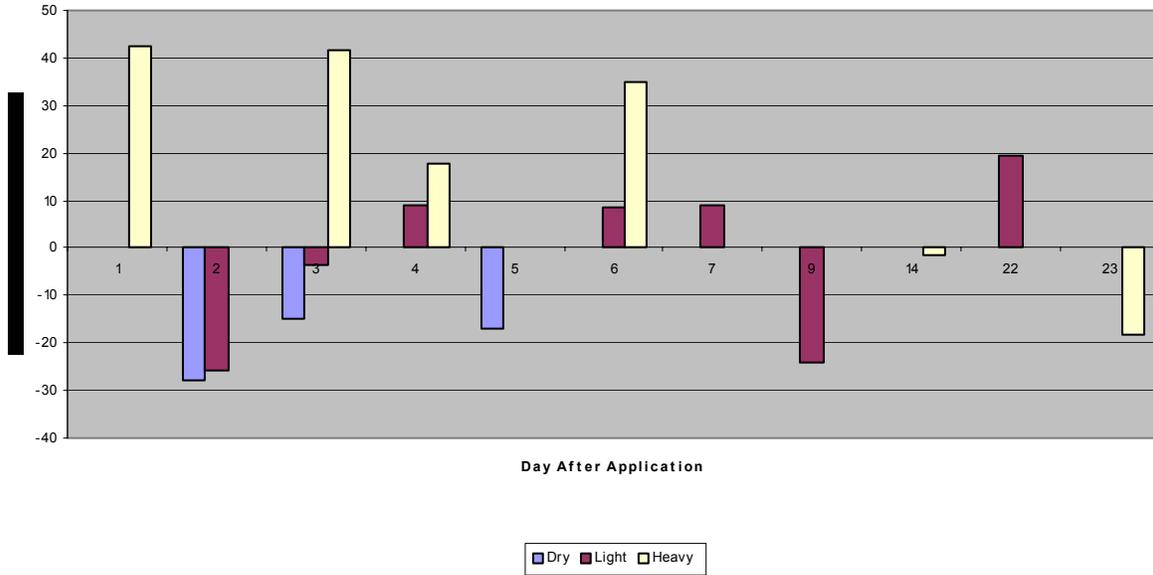


Fig. B Change in Dermal TC by moisture level during monitoring

