**New Insights from a Study of Chlorpyrifos Toxicology in Humans**

Seth G. McGrew

A thesis

submitted in partial fulfillment of the

requirements for the degree of

Master of Science

University of Washington

2018

Committee:

Elizabeth A. (Lianne) Sheppard

Richard A. Fenske

Program Authorized to Offer Degree:

Environmental and Occupational Health Sciences

University of Washington

**Abstract**

New Insights from a Study of Chlorpyrifos Toxicology in Humans

Seth G. McGrew

Chair of the Supervisory Committee:

Professor Elizabeth A. (Lianne) Sheppard, PhD, MSc

Department of Environmental and Occupational Health Sciences

**Background:**

Chlorpyrifos is among the most heavily used conventional insecticides in the United States. The public is chronically exposed to low levels of chlorpyrifos in food crop residues and drinking water. In 1971, Dow Chemical Company conducted a study in human subjects to determine a level of chlorpyrifos that can be ingested every day without inhibiting plasma cholinesterase (plasma-ChE) and red blood cell cholinesterase (RBC-ChE), two markers of the acute toxicological effect of organophosphate exposure. The study's only reported significant effect was plasma ChE inhibition in the subjects treated at 100μg/kg/day, the highest dose level tested.

**Objective:**

The primary goal of this thesis is to reassess the original safety evaluation to see whether new analysis supports revision of the previously reported effect of treatment on plasma ChE activity. The project also seeks to characterize the recovery trend that follows treatment. A review of regulatory documents intends to clarify past use of Dow's safety evaluation in EPA's regulatory process.

**Methods:**

The analysis uses repeated measures ANOVA and linear mixed effects (LMM) regression to examine plasma ChE activity data for treatment effects at each dose level. Trends in recovery of plasma ChE activity are assessed by LMM regression with splines for subjects treated at 30µg/kg/day and 100µg/kg/day.

**Results:**

The original analysis by repeated measures ANOVA used a subset of data for the 30µg/kg/day group to make its determination of no treatment effect on plasma ChE activity ($p = 0.12$). When all measurements are included in the ANOVA, the data indicate a significant effect at that dose level ($p = 0.012$). An LMM regression model that imposes no pattern of dose dependence estimates significant inhibitory effects on plasma ChE at the 30µg/kg/day and 100µg/kg/day levels ($p < 0.0001$), and nearly significant inhibition at 14µg/kg/day ($p = 0.066$). Inhibitory effect estimates are -0.029, -0.070, and -0.349µmoles acetate/min/ml/day for treatment at 14, 30, and 100µg/kg/day, respectively. LMM regression that assumes a linear relationship between dose and treatment effect estimates significant inhibition for treatment at all three dose levels. LMM regression with splines estimates that subjects treated at 30µg/kg/day recovered their plasma ChE activity at 0.047µmoles acetate/min/ml/day, and subjects treated at 100µg/kg/day recovered more than twice as rapidly at 0.101µmoles acetate/min/ml/day.

**Conclusions:**

The original analysis failed to detect an effect from treatment at 30µg/kg/day that is well supported by the data. Plasma ChE inhibition by chlorpyrifos has approximately linear dose dependence for repeated daily treatments at 14, 30, and 100µg/kg/day. Acute plasma ChE inhibition may precipitate more rapid recovery than less severe inhibition. Earlier identification of the treatment effect at 30µg/kg/day would have reduced the maximum acceptable oral dose established by the EPA in 1986.

**Acknowledgements**

I would like to thank my committee members for their assistance with this thesis:

Lianne provided the idea for the project and has contributed countless hours of expert advice and encouragement. Throughout, she has taken care to emphasize the importance of responsible scientific interpretation and reporting. Many thanks.

I am very grateful to Richard for sharing his expertise in regulatory matters pertaining to chlorpyrifos. His input helped to identify critical reference materials and clarify the reasoning behind EPA's sometimes confusing regulatory decisions.

In addition, I would like to thank Professor Terrance Kavanagh for his comments on the cholinesterase activity assay, and Professor Lucio Costa for his suggestions regarding the apparent pattern of recovery from cholinesterase inhibition.

This project was supported in part by a generous grant from the Boeing Company to the University of Washington's Department of Environmental and Occupational Health Sciences.

## Table of Contents

**List of Tables**

**List of Figures**

# 1. Introduction

## 1.1 Chlorpyrifos: the chemical of interest

The broad purpose of this research is to more fully understand human health effects associated with exposure to chlorpyrifos, an organophosphate pesticide originally formulated and manufactured by Dow Chemical Company in 1965. Initially designated DOWCO 179, chlorpyrifos has been marketed under the trade names Dursban, Empire 20, Equity, Whitmire PT 270, and Lorsban, among others (US EPA, 2006). It is a broad-spectrum insecticide, acaricide, and miticide still routinely applied to nearly 50 crops including corn, alfalfa, soybean, almond, orange, cotton, and certain vegetable crops. According to its manufacturer, chlorpyrifos is registered for use in nearly 100 countries (Dow, 2017). The US Environmental Protection Agency (EPA) has called it the most-used conventional insecticide in the United States, with an estimated 6 million pounds applied annually to 10 million acres of US cropland between 2009 and 2013 (US EPA, 2016b). It is only moderately persistent in water and soil, and has low potential to bioaccumulate in animals (Christensen, 2009). Dow continues to tout the pesticide's efficacy, low cost, minimal impact on beneficial insects, and ease of integration into existing pest- and resistance management programs (Dow, 2017). Nonetheless, because of its prevalent use in agriculture, chlorpyrifos has considerable potential to impact human health. In recent decades, EPA has found cause to restrict or suspend several uses of chlorpyrifos, citing concern over its potential adverse impacts to humans and other non-target organisms (US EPA, 2016b).

Chlorpyrifos is sold as a liquid concentrate or crystalline solid of low aqueous solubility. Prior to use, it is dissolved in a non-polar solvent, then further diluted in water to a typical final concentration of 0.025% to 0.05% for direct spray application to crops. Pesticide mixers, pest

control operators, and agricultural workers are the occupational groups most likely to be affected by acute chlorpyrifos poisoning. Their predominant routes of exposure are dermal absorption and inhalation (US EPA, 2016a). Since EPA's curtailment of residential uses in the early 2000's, non-occupational exposures to chlorpyrifos result primarily from dietary intake of residue in food and drinking water. Incidental exposures by inhalation, ingestion, or dermal absorption may also occur near agricultural operations. EPA now considers the acute and chronic health risk from non-occupational exposures to be acceptably low; however, epidemiologic studies begun in the late 1990's and early 2000's have raised concern that typical exposures during gestation and early childhood are associated with developmental deficits in later childhood (Whyatt, 2004; Rauh, 2006; Rauh, 2011; US EPA 2016a). Since 2007, prominent environmental advocacy groups have maintained an active petition for EPA to protect children's health by canceling all tolerances for chlorpyrifos (NRDC/PANNA, 2007).

### 1.2 Cholinergic properties in humans

Like other organophosphate pesticides, including its structural analogs parathion and diazinon, chlorpyrifos disrupts neural impulse transmission in target insects and arachnids. After an insect is exposed, initial biotransformation steps catalyzed by cytochrome p450 (CYP) enzymes convert chlorpyrifos to both a metabolically active oxon form and harmless breakdown products. When the oxon metabolite binds and inhibits the enzyme acetylcholinesterase (AChE) in neuromuscular synapses, it slows the removal of acetylcholine (ACh) that typically follows a neurotransmission event. Accumulated ACh molecules signal downstream motor neurons to fire repeatedly, leading to uncoordinated action and eventual depletion of energy stores in affected muscle cells. This results in paralysis and death of the animal (Casarett, 2010).

Because many CYP-catalyzed biotransformation processes and some ACh-mediated neurotransmission functions are conserved in vertebrates, organophosphate pesticides have cholinergic effects in humans and other higher animals as well as target species. The acute reaction to chlorpyrifos poisoning in humans is characterized by weakness, headache, tremor, sweating, and salivation. In severe cases, victims enter a state of flaccid paralysis and may die from respiratory failure or cardiac arrest. Cholinergic effects, typically quantified as percentages of normal cholinesterase (ChE) activity, are detectable in ChE isoforms found outside of the nervous system as well, including butyryl cholinesterase in blood plasma (plasma ChE) and acetylcholinesterase in erythrocytes (RBC-ChE). Though its normal physiologic role is unknown, plasma ChE responds more readily to chlorpyrifos inhibition than AChE in the central nervous system, and its activity can be a useful marker of organophosphate exposure (Stefanidou, 2009; Tang, 2001).

### 1.3 The DOWCO 179 safety evaluation study

EPA's current regime for pesticide risk assessment uses animal studies and physiologically based pharmacokinetic modeling (PBPK) to estimate human toxicological effects and levels of exposure. Prior to 1998, though, EPA risk assessment for chlorpyrifos relied heavily on one of the few existing studies to directly measure toxicological effects of dietary chlorpyrifos in humans. The research is described in Dow Chemical Company's unpublished 1972 report, "Safety Evaluation of DOWCO 179 in Human Volunteers," by Coulston, Golberg, and Griffin (Coulston, 1972). The study considered inhibition of RBC-ChE and plasma ChE as indicators of cholinergic response, and reported a threshold level for daily ingestion of chlorpyrifos below which there was no significant depression of plasma ChE activity in healthy adults.

The specific question addressed in the Dow study asked whether daily ingestion of chlorpyrifos reduces plasma ChE or RBC-ChE activity over time in adult males. Its design randomized 16 healthy adult volunteers into four treatment groups. Once per day during the treatment phase, each subject ingested a tablet containing either a placebo or crystalline chlorpyrifos at one of three assigned dose levels. Twice per week during treatment and recovery periods, investigators collected a blood sample from each subject. Samples were immediately analyzed for plasma ChE and RBC-ChE activity. The study featured a 9- to 27-day treatment period and twice-weekly follow up until all treated subjects' plasma ChE and RBC-ChE activities returned to baseline. Investigators then used repeated measures ANOVA to make pairwise assessment of the probability that treatment-day-specific cholinesterase activity differed between each treatment group and the untreated control group. Coulston reported that analysis by repeated-measures ANOVA found a significant difference in plasma ChE activity for only those subjects treated at the highest dose of 100μg/kg/day, and no significant difference in RBC-ChE activity at any dose level.

The current research reevaluates the DOWCO 179 treatment study, noting its design limitations and analyzing its dataset using both the original repeated-measures ANOVA approach and more modern computational regression methods. This work also includes an analysis of the previously unexplored association between dose level and recovery to normal plasma ChE activity following exposure. A final discussion considers the DOWCO 179 study's historic role in EPA's regulatory process, and its relevance to modern understanding of chlorpyrifos toxicology.

### 1.4 Regulatory status over time

Since its introduction in 1965, EPA has approved the sale, distribution and use of chlorpyrifos through uninterrupted registration under the Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA). In 1996, the Food Quality Protection Act (FQPA) amended FIFRA to require that EPA establish a reference level for pesticide residue exposure below which there is reasonable certainty of no harm. The amendment called for reassessment of all pesticide tolerances currently in effect, and required EPA to develop more refined risk assessment strategies to estimate non-occupational risks from drinking water, residential exposures, aggregate exposure, and exposure to multiple pesticides with a common mechanism of toxicity. Notably, FQPA instructed EPA to adjust its tolerances, as warranted, to account for possible higher pesticide sensitivity among infants and children. In compliance with FQPA, EPA eliminated or phased out most residential uses of chlorpyrifos beginning in 2000, with exceptions for insect baits encased in child-resistant packaging. Approval for the use of products containing chlorpyrifos on tomatoes, post-bloom apples, and non-dormant grape vines was also discontinued. Later, in 2002, EPA reduced the allowable frequency and/or number of applications to citrus fruit, alfalfa, corn, cotton, peanuts, and tree nuts, among other crops (US EPA, 2006).

As required under FQPA, EPA now renews pesticide registrations according to a 15-year cycle. To secure a chemical's eligibility for reregistration, a manufacturer must submit evidence to demonstrate the product's continued relevance and efficacy. In addition, EPA conducts extensive assessments of the potential health risks and ecological impacts associated with expected use of the chemical. EPA most recently approved chlorpyrifos for reregistration in

2006, with the list of allowable uses reflecting the restrictions imposed in 2000 and 2002. Since that time, the chemical has been under review for the upcoming reregistration period. In 2012, as a condition for future reregistration eligibility, the Agency imposed further interim restrictions on chlorpyrifos use to reduce non-occupational risks. The new rules limit application rates and create buffer zones around croplands near sensitive sites such as schools, daycare facilities, homes, parks, and hospitals (US EPA, 2016b).

In 2007, shortly after Dow had secured its current FIFRA registration, Pesticide Action Network of North America (PANNA) and the Natural Resources Defense Council (NRDC) delivered a petition urging EPA to suspend all tolerances for chlorpyrifos (NRDC/PANNA, 2007). In support of their position, the petitioners cited epidemiological evidence that occupational exposures among farm workers and typical residues on food could adversely impact cognitive development in children and developing fetuses. In addition, they contended that aggregate exposures among residents of agricultural areas with high pesticide loading in watersheds could exceed the current limit for ingestion specified under the Federal Food, Drug, and Cosmetics Act (FFDCA). In response to the petition, EPA assembled a scientific advisory panel (SAP) to integrate the results of toxicity studies with epidemiological data, and to specifically consider the potential for long-term neurodevelopmental effects with early life exposure. Between 2008 and 2012, the SAP convened five times, but was unable to deliver a recommendation, based on the existing research, to either approve or deny the proposed revocation. Ultimately, PANNA and NRDC secured orders through the 9th District Court of Appeals requiring EPA to expedite its risk assessment and rule on the petition by April, 2017. On

March 29, 2017, newly appointed EPA Administrator Scott Pruitt formally denied the request for revocation of tolerances (US EPA, 2017).

### 1.5 Goal of this thesis

The goal of this work is to develop new insights from the 1972 Coulston study of chlorpyrifos toxicology in adult men. First, a review of the study's documentation informs a critique of the original study design and its implementation. In attempting to reproduce the reported results by the original method, we hope to understand decisions made in the first analysis. Second, we determine whether analytical methods that use the data set more efficiently may show evidence of a treatment effect that the original method was unable to detect. Third, by characterizing the subjects' patterns of recovery from treatment, we hope to better understand the process by which cholinesterase activity is restored following organophosphate exposure. Finally, by reviewing past regulatory documents, we intend to clarify the Coulston study's historic role in chlorpyrifos regulation and explain the reasons for its discontinued use as a basis for determinations of allowable exposure.

## 2. Determination of treatment effect for DOWCO 179

### 2.1 Overview of the DOWCO 179 safety evaluation

Albany Medical College's Institute of Experimental Pathology and Toxicology oversaw the safety evaluation for DOWCO 179 during the summer of 1971. The methods and results are described in the unpublished report by Coulston, Golberg, and Griffin, "Safety Evaluation of DOWCO 179 in Human Volunteers" (Coulston, 1972). The study included 16 healthy adult male inmates from a pool of volunteers at the Clinton Correctional Facility, a maximum-security prison located in Dannemora, New York. Subjects were admitted to the facility's hospital for the duration of the study, but were permitted to leave while performing daily chores and duties. Prior to enrollment in the study, each subject received a physical examination to confirm his normal clinical status; testing included chest X-ray, electrocardiogram, urinalysis, and assays of hematology and serum chemistry. The 16 chosen volunteers were randomized into four experimental groups, one group acting as a negative control, and the others assigned to receive daily DOWCO 179 treatment at one of three levels: 14µg/kg, 30µg/kg, or 100µg/kg of body weight. Because Coulston's report does not specify an initial study design, the planned duration of treatment and exact criteria for terminating the study are unknown.

During the treatment period, subjects received a daily administration of either placebo or DOWCO 179 in tablet form with the breakfast meal. Treatments continued for 48, 27, 21 and 9 days at the control, 14µg/kg, 30µg/kg, and 100µg/kg levels, respectively. Twice weekly, heparinized blood samples were taken and separately analyzed for RBC-ChE and plasma ChE activity using the method described by Nabb and Whitfield (Nabb, 1967). The method employs automated pH stat titration to quantify acetic acid evolved when the enzyme substrate,

acetylcholine perchlorate, is hydrolyzed by cholinesterase. Activity was measured in units of μmoles acetate/min/ml plasma and recorded to the nearest tenth. For the duration of treatment, subjects were monitored daily for general symptoms of declining health. Additional blood samples, collected weekly, were used for hematologic counts and automated assessments of serum chemistry. Weekly urinalyses included observations of turbidity, pH, specific gravity, and cellular content. Samples of urine were also retained for later assay to detect DOWCO 179 and its biotransformation products.

Following completion of the treatment phase, monitoring of cholinesterase activity continued semiweekly until values returned to baseline levels. Samples retained for blood counts and urinalyses were collected once per week during recovery. At the study's conclusion, hematologic analyses found no evidence of a treatment effect on packed cell volumes, hemoglobin concentration, total white blood cell counts, or differential white blood cell counts for any subject (Coulston, 1972). Similarly, no changes in serum chemistry parameters were associated with DOWCO 179 ingestion at any dose level. The report does not include the urinalysis data, but states that routine urinalyses noted no change with treatment, and subsequent chemical analysis of stored urine samples failed to detect DOWCO 179, its oxon analog, or its decomposition product 3,5,6-trichloro-2-pyridinol (TCPy). Limits of detection for the assays are not mentioned. Coulston did report one instance of possible significance: on the ninth day of treatment at 100μg/kg, one subject complained of a runny nose, dizziness, and faintness. After examining the patient, investigators attributed the symptoms to a cold; nonetheless, they immediately discontinued treatment in the 100μg/kg dose group.

Coulston reported no statistically significant difference in plasma ChE activity between the control group and subjects treated for 28 days at 14µg/kg. In the dose group treated for 21 days at 30µg/kg, average plasma ChE activity dropped to 70% of baseline levels. When compared on a day-by-day basis for all time points during treatment, group plasma ChE levels averaged 87% of concurrent control group levels. However, because the difference was not found to be statistically significant, the investigators judged 30µg/kg to be a threshold level in that daily treatment for 21 days showed evidence of exposure, but did not have a significant toxicological effect in the human subjects. The study's most notable outcome was a statistically significant decline in plasma ChE activity for subjects treated at 100mg/kg. In that group, after 9 days of treatment, plasma ChE levels averaged 34% of baseline. When compared on a day-by-day basis, group plasma ChE levels averaged 70% of concurrent control group levels during treatment. The precipitous drop in activity led the researchers to discontinue treatment in the 100µg/kg group after the ninth day.

The study also found that no level of treatment produced a detectable change in RBC-ChE activity relative to controls. The text explains that, though no change in RBC-ChE was detected at the higher dose levels, a significant difference *was* noted for the 14µg/kg subject group. The researchers dismissed that finding as an artifact, however, presumably on the assumption that any true dose-dependent response seen at the lowest dose level would be more pronounced, rather than undetectable, at higher doses. Because RBC-ChE was not observed to respond to treatment at the higher dose levels, the current analysis focuses entirely on the activity of plasma ChE.

## 2.2 Original analysis and limitations imposed by realized study design

The DOWCO 179 study used a repeated measures design. That is, successive measurements of plasma cholinesterase activity were noted for each subject over time, at semiweekly (two-to-four-day) intervals. Due to natural heterogeneity within the population, individual baseline measurements and response to a treatment or disease can vary considerably from person to person (Fitzmaurice, 2004). An advantage to the repeated measures design is that it separately tracks each individual's response to treatment over time, rather than relying on overall estimates of response to treatment among potentially heterogeneous individuals. This separation by identity allows for the errors due to between-subject variation in response to be excluded from the estimated mean response within a dose group, yielding a dose-specific treatment effect estimate that is less influenced by intrinsic between-subject differences than would be possible using a cross-sectional approach.

Details of the safety evaluation's statistical analysis appear in an unpublished report, "Analysis of Plasma and Red Blood Cell-Cholinesterase Levels in Humans Ingesting DOWCO 179," by Colin Park of Dow Chemical Company's Mathematical Applications Division (Park, 1972). In his assessment, Park observes that day-to-day variation in ChE activity for a given individual are large, but that day-to-day trends are consistent from person to person. This fact precludes a simple assessment of day-to-day changes in one treatment group without consideration of parallel variation in at least one control group. Further, Park notes, there is considerable variation in baseline CHE levels between subjects. This implies that baseline differences cannot be ignored in comparisons of plasma ChE activity between subject groupings. Finally, Park notes, because the repeated measurements on a given subject are correlated with

one another, the study data violate the assumption of independence required for valid analysis by simple two-way ANOVA or least squares regression.

To account for the lack of independence and subject-to-subject variation in his test for a difference in response between groups, Park uses the method of repeated measures ANOVA, a mixed model analysis of variance strategy with random subjects and fixed day and dose variables. Repeated measures ANOVA applies an F-test to the ratio computed as the mean-squared error due to mean differences between dose groups, divided by an expression of the residual mean squared error for the model that excludes error introduced by natural variation between subjects. Though modern computer software requires only a fraction of a second to complete an ANOVA of this size, in 1972 the error calculations were made laboriously by hand, with the p-value read from an F-distribution table. The null hypothesis for the ANOVA assumes that there is a constant mean difference between treatment and control over time in the interaction of dose-by-day. The alternative hypothesis is a general alternative that allows for any kind of difference between treatment and control in the dose-by-day interaction.

While repeated measures ANOVA appropriately managed the lack of independence in longitudinal data based on the technology available at the time, it also imposed a requirement that was not well satisfied by the realized design of the Dowco study. To be included in the ANOVA's sum-of-squares error calculations, the data needed to be balanced. This means that every data point from one group in a comparison must be balanced by a measurement taken at the same time point for every other group in the comparison. For this study that meant that on every day included in a comparison, all dose groups had to show a measurement for each subject.

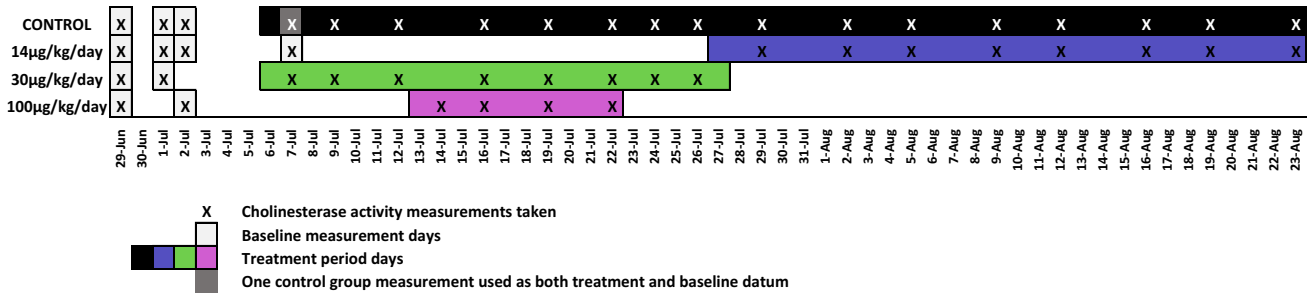
For reasons not clarified in the text, treatment periods in the study did not start on the same day for all dose groups. Furthermore, the duration of treatment differed between dose groups; for example, there were 8 treatment day measurements between July 29th and August 23rd for the 14µg/kg subjects, 8 treatment day measurements between July 7th and July 26th for the 30µg/kg subjects, and 4 treatment day measurements between July 14th and July 22nd for the 100µg/kg subjects. Two graphic illustrations of the mismatch in treatment timing appear in Figures 2.1 and 2.2. Because the measurements were not balanced in time among the three dose groups and the control group, it was not possible to compare all dose groups together in a single ANOVA. That is, with the statistical tools available in the 1970's, the realized study design permitted only one-at-a-time comparisons between each treated dose group and the control. Further, each pairwise comparison required some unbalanced control group data to be omitted; in Table 2.1 below, all control group measurements not highlighted in green were excluded from comparisons in which they were not balanced by treatment group measurements. In the case of the 100µg/kg group, *treatment* data from July 14th had to be discarded because no control data were collected on that day; the omission left only three treatment time points and two baseline time points in the ANOVA comparison between 100µg/kg subjects and controls.

Table 2.1 Plasma ChE time point measurements included in Park's three original comparisons by repeated measures ANOVA: 14µg/kg/day subjects vs. controls, 30µg/kg/day subjects vs. controls, and 100µg/kg/day subjects vs. controls

| day of measurement | | control subjects | | | | 14µg subjects | | | | control subjects | | | | 30µg subjects | | | | control subjects | | | | 100µg subjects | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **date** | **day** | **2** | **4** | **9** | **16** | **7** | **8** | **14** | **15** | **2** | **4** | **9** | **16** | **10** | **11** | **12** | **13** | **2** | **4** | **9** | **16** | **1** | **3** | **5** | **6** |
| 29-Jun | 0 | 5.1 | 5.1 | 3.1 | 4.0 | 4.1 | 3.3 | | | 5.1 | 5.1 | 3.1 | 4.0 | 4.0 | 3.5 | 7.0 | 3.6 | 5.1 | 5.1 | 3.1 | 4.0 | 4.0 | 5.2 | 5.2 | 4.1 |
| 1-Jul | 2 | | | | 3.2 | | | 4.0 | 4.8 | | | | 3.2 | 4.1 | 3.1 | 6.9 | 5.3 | | | | 3.2 | | | | |
| 2-Jul | 3 | 4.8 | 4.9 | 3.0 | | 4.0 | 3.2 | | | 4.8 | 4.9 | 3.0 | | | | | | 4.8 | 4.9 | 3.0 | | 4.1 | 4.8 | 5.7 | 4.8 |
| 7-Jul | 8 | 4.6 | 4.6 | 5.0 | 4.2 | | | 5.4 | 4.7 | 4.6 | 4.6 | 5.0 | 4.2 | 4.1 | 3.1 | 6.8 | 3.8 | 4.6 | 4.6 | 5.0 | 4.2 | | | | |
| 9-Jul | 10 | 6.1 | 4.1 | 3.1 | 4.6 | | | | | 6.1 | 4.1 | 3.1 | 4.6 | 3.9 | 3.1 | 6.9 | 3.3 | 6.1 | 4.1 | 3.1 | 4.6 | | | | |
| 12-Jul | 13 | 5.4 | 5.3 | 3.3 | 4.3 | | | | | 5.4 | 5.3 | 3.3 | 4.3 | 4.4 | 2.8 | 5.7 | 2.9 | 5.4 | 5.3 | 3.3 | 4.3 | | | | |
| 14-Jul | 15 | | | | | | | | | | | | | | | | | | | | | 3.0 | 5.2 | 6.1 | 3.0 |
| 16-Jul | 17 | 5.2 | 6.5 | 3.6 | 4.5 | | | | | 5.2 | 6.5 | 3.6 | 4.5 | 5.2 | 3.0 | 7.8 | 2.7 | 5.2 | 6.5 | 3.6 | 4.5 | 4.1 | 4.6 | 5.9 | 2.6 |
| 19-Jul | 20 | 4.9 | 5.4 | 3.5 | 4.5 | | | | | 4.9 | 5.4 | 3.5 | 4.5 | 4.3 | 2.7 | 6.2 | 2.7 | 4.9 | 5.4 | 3.5 | 4.5 | 1.5 | 2.8 | 4.6 | 1.7 |
| 22-Jul | 23 | 4.3 | 4.8 | 3.4 | 4.1 | | | | | 4.3 | 4.8 | 3.4 | 4.1 | 3.3 | 2.2 | 5.4 | 2.1 | 4.3 | 4.8 | 3.4 | 4.1 | 0.9 | 1.6 | 3.5 | 0.8 |
| 24-Jul | 25 | 4.7 | 4.6 | 3.3 | 4.5 | | | | | 4.7 | 4.6 | 3.3 | 4.5 | 3.4 | 2.3 | 4.6 | 2.1 | 4.7 | 4.6 | 3.3 | 4.5 | 1.0 | 1.7 | 3.9 | 1.4 |
| 26-Jul | 27 | 4.8 | 5.2 | 3.4 | 4.1 | | | | | 4.8 | 5.2 | 3.4 | 4.1 | 3.4 | 2.3 | 5.6 | 2.2 | 4.8 | 5.2 | 3.4 | 4.1 | 1.5 | 2.2 | 4.3 | 1.4 |
| 29-Jul | 30 | 4.8 | 4.8 | 3.3 | 3.1 | 4.4 | 3.1 | 5.5 | 4.8 | 4.8 | 4.8 | 3.3 | 3.1 | 3.5 | 2.5 | 5.9 | 2.4 | 4.8 | 4.8 | 3.3 | 3.1 | 1.7 | 2.8 | 4.8 | 1.8 |
| 2-Aug | 34 | 4.8 | 5.1 | 3.2 | 4.4 | 4.7 | 3.0 | 5.5 | 4.7 | 4.8 | 5.1 | 3.2 | 4.4 | 3.9 | 2.8 | 4.8 | 2.7 | 4.8 | 5.1 | 3.2 | 4.4 | 2.4 | 3.7 | 5.4 | 2.4 |
| 5-Aug | 37 | 4.5 | 4.5 | 2.6 | 4.1 | 3.9 | 2.6 | 5.1 | 3.5 | 4.5 | 4.5 | 2.6 | 4.1 | 3.5 | 2.6 | 6.3 | 2.1 | 4.5 | 4.5 | 2.6 | 4.1 | 2.3 | 3.3 | 4.7 | 2.3 |
| 9-Aug | 41 | 4.4 | 4.1 | 3.2 | 3.5 | 3.5 | 2.7 | 3.3 | 3.8 | 4.4 | 4.1 | 3.2 | 3.5 | 3.3 | 2.5 | 6.3 | 1.5 | 4.4 | 4.1 | 3.2 | 3.5 | 2.3 | 3.6 | 3.7 | 2.0 |
| 12-Aug | 44 | 3.9 | 4.4 | 3.0 | 3.8 | 3.6 | 2.0 | 4.8 | 3.4 | 3.9 | 4.4 | 3.0 | 3.8 | | | | | 3.9 | 4.4 | 3.0 | 3.8 | 2.6 | 3.6 | 5.4 | 2.6 |
| 16-Aug | 48 | 4.3 | 4.8 | 2.8 | 4.1 | 3.8 | 2.6 | 5.2 | 3.9 | 4.3 | 4.8 | 2.8 | 4.1 | 4.0 | 2.9 | 6.7 | 4.1 | 4.3 | 4.8 | 2.8 | 4.1 | 4.1 | 4.5 | 6.7 | 3.4 |
| 19-Aug | 51 | 4.7 | 4.8 | 3.5 | 5.2 | 3.5 | 2.8 | 5.7 | 4.0 | 4.7 | 4.8 | 3.5 | 5.2 | 4.1 | 3.1 | 6.7 | 3.5 | 4.7 | 4.8 | 3.5 | 5.2 | 3.6 | 5.4 | 6.5 | 4.0 |
| 23-Aug | 55 | 4.5 | 5.4 | 3.2 | 4.4 | 3.4 | 2.9 | 5.7 | 3.8 | 4.5 | 5.4 | 3.2 | 4.4 | 4.3 | 2.9 | 6.5 | 3.2 | 4.5 | 5.4 | 3.2 | 4.4 | 3.5 | 4.8 | 6.4 | 3.7 |
| 26-Aug | 58 | 4.8 | 5.4 | 3.2 | 5.1 | 4.7 | 3.2 | 6.1 | 5.1 | 4.8 | 5.4 | 3.2 | 5.1 | 4.7 | 3.1 | 7.6 | 3.3 | 4.8 | 5.4 | 3.2 | 5.1 | 3.8 | 5.5 | 5.9 | 3.7 |
| 30-Aug | 62 | 5.0 | 5.0 | 3.3 | 4.7 | 4.7 | 3.7 | 6.0 | 4.6 | 5.0 | 5.0 | 3.3 | 4.7 | | | | | 5.0 | 5.0 | 3.3 | 4.7 | | | | |

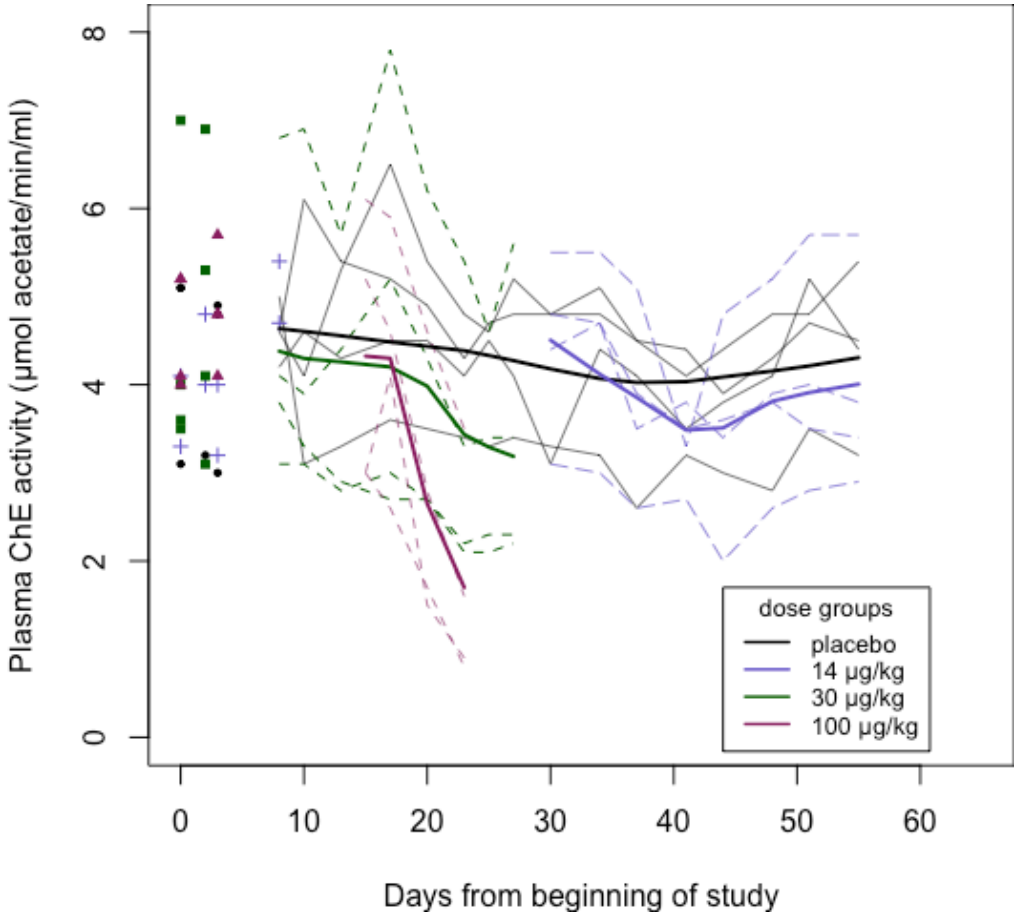
first baseline measurement
second baseline measurement
baseline measurements omitted from 30µg group ANOVA
treatment measurements
recovery measurements (not included in analysis)

Figure 2.1 Timing of treatment periods and ChE activity measurements by dose group




X Cholinesterase activity measurements taken
Baseline measurement days
Treatment period days
One control group measurement used as both treatment and baseline datum

Figure 2.2 Unadjusted plasma ChE measurements for baseline days and treatment period: individual data and smoothed group means vs. day of study




The pairwise approach effectively reduced the number of samples included in the ANOVA, resulting in error estimates that were noisier than necessary. Even within the paired data subsets, lack of balance required investigators to exclude some valid measurements from the analysis. Both of these actions reduced the study's power to detect differences with treatment. Finally, misalignment in the timing of treatments made tests for differing inhibitory effects among the non-zero dose groups impossible.

## 2.3 Replication of original treatment analysis by repeated measures ANOVA

This and all subsequent analyses used RStudio Version 1.0.136. Script code is included in Appendix E.

Attempts to replicate the reported results by repeated measures ANOVA were successful for the two higher dose categories, but required some recoding of dates to balance baseline data (See Table 2.1 for successful date coding). When the data are recoded to align all baseline measurements on a common day 0, Park's results for the 100µg/kg group are exactly replicable (see Table 2.2 for summarized results and Appendix B for ANOVA output). In the 30µg/kg group comparison, Park excluded all data for the second baseline day from the analysis; the result of this omission is that the difference in mean plasma ChE activity between the treatment and control groups is not significant, with p = 0.12. If the second day's baseline data are included, the difference is quite significant at p = 0.012. Results for the 14µg/kg dose group were only replicable when control subject #16's first baseline datum was ignored and his second baseline day and first placebo day were designated baseline days 1 and 2, respectively.

Table 2.2 Original and updated assessments of treatment effect by repeated measures ANOVA

| **Dose (µg/kg/day)** | ***Original Analysis*** | | | ***Replication of Original Analysis*** | | |
|---|---|---|---|---|---|---|
| | n | Degrees of Freedom* | p-value | n | Degrees of Freedom* | p-value |
| 14 | 80 | 9 | 0.41 | 80 | 9 | 0.41 |
| 30 | 72 | 8 | 0.12 | 80 | 9 | 0.012 |
| 100 | 48 | 4 | <0.001 | 48 | 4 | <0.001 |

* degrees of freedom for day*dose parameter = number of treatment days - 1

## 2.4 New treatment analysis by linear mixed effects regression

Like repeated measures ANOVA, linear mixed effects (LMM) regression compensates for the lack of independence between repeated measurements on the same individual by separating within-subject variance components from the model's residual variance expression. In essence, the repeated measures ANOVA is also a form of least-squares mixed effects model with all categorical variables; Model 0 below expresses the ANOVA from the preceding section as a linear mixed effects model with categorical representations of dose, treatment day, and subject. Unlike repeated measures ANOVA, though, modern LMM regression, first proposed in the early 1980's by Laird and Ware, uses a computational algorithm to estimate regression coefficients by maximum likelihood (Laird, 1982). This method of estimation, because it tolerates data imbalance, is better suited to the temporally misaligned DOWCO 179 study data than the sum-of-squares method used in repeated measures ANOVA. Using all of the study's baseline and treatment data points in a single model, the LMM regression can simultaneously estimate treatment effects for all three dose levels and controls.

**Model 0: all-categorical ANOVA model**

$$
\begin{aligned}
PChE_{ijt} \;=\;& \boldsymbol{\beta}_{0j} \times \mathbf{I}(\text{dose} = j) && \textit{intercept = baseline mean}\\
+\;& \boldsymbol{\beta}_{1j} \times \mathbf{I}(\text{dose} = j) \times t_{trt} && \textit{treatment effect}\\
+\;& \mathbf{b}_i && \textit{subject random effect}\\
+\;& \varepsilon_{ijt} && \textit{residual noise}
\end{aligned}
$$

$$b_i \sim \text{iid } N(0, \sigma^2_b), \quad \varepsilon_{ijt} \sim \text{iid } N(0, \sigma^2_{\varepsilon})$$

- $PChE_{ijt}$ is the plasma cholinesterase activity level in μmoles acetate/min/ml for subject i in dose group j on day t of the study.
- Dose j is categorical, identifying the dose groups 0, 14, 30, and 100μg/kg/day.
- Time $t_{trt}$ is a categorical variable: all baseline and treatment days are weighted equally in the model.

A complication that the DOWCO 179 study design introduces to the multi-dose model is potential confounding of the treatment effect by day-to-day variation in laboratory results. Because data were collected on different days for different dose groups, a day-to-day trend unrelated to the treatment response could potentially mask or exaggerate a true treatment effect. Such a trend is apparent in the smoothed plot of mean control measurements in Figure 2.2. When present, though, this confounding can be corrected in the model by including 'day of study' as a categorical predictor with a level for each day; the computed coefficients represent the contribution of mean date-specific variations to the modeled plasma ChE level.

Our LMM treatment models employ two different indices of time: the categorical 'day of study' index (t) in the term predicting the day-to-day variability, and a continuous count of treatment days ($t_{trt}$) used in the dose*days expression. Coulston's report did not specify the exact timing of blood sampling relative to treatment; however, Park's analysis did include notes that specified the starting and ending dates for each dose group's treatment period, along with handwritten treatment day indices that corroborated those notes (See Appendix A). Because Park considered the day after the first administration to be treatment day 1, it appears that each measurement represents a subject's ChE activity level one day after the most recent administration. For every treatment period measurement, we have assumed, in the absence of more detailed information, that approximately 24 hours elapsed between the preceding treatment and the blood sample collection for that measurement. For each measurement date in the study calendar, Table 2.3 shows the corresponding t and $t_{trt}$ that were used in the regression model.

Table 2.3 Time indices in the LMM regression model of treatment effect

| date | t | $t_{trt}$ | | | |
|---|---|---|---|---|---|
| | | control group | 14µg group | 30µg group | 100µg group |
| 29-Jun | 0 | 0 | 0 | 0 | 0 |
| 1-Jul | 2 | 0 | 0 | 0 | |
| 2-Jul | 3 | 0 | 0 | | 0 |
| 7-Jul | 8 | 1 | 0 | 1 | |
| 9-Jul | 10 | 3 | | 3 | |
| 12-Jul | 13 | 6 | | 6 | |
| 14-Jul | 15 | | | | 1 |
| 16-Jul | 17 | 10 | | 10 | 3 |
| 19-Jul | 20 | 13 | | 13 | 6 |
| 22-Jul | 23 | 16 | | 16 | 9 |
| 24-Jul | 25 | 18 | | 18 | |
| 26-Jul | 27 | 20 | | 20 | |
| 29-Jul | 30 | 23 | 1 | | |
| 2-Aug | 34 | 27 | 5 | | |
| 5-Aug | 37 | 30 | 8 | | |
| 9-Aug | 41 | 34 | 12 | | |
| 12-Aug | 44 | 37 | 15 | | |
| 16-Aug | 48 | 41 | 19 | | |
| 19-Aug | 51 | 44 | 22 | | |
| 23-Aug | 55 | 48 | 26 | | |
| 26-Aug | 58 | | | | |
| 30-Aug | 62 | | | | |

baseline day
treatment day
recovery day

**t** day of study
$t_{trt}$ treatment day index

Our analysis employed two versions of a treatment model that differ only in their coding of the dose level (See Models 1 and 2 below). Because Model 1 considers dose as a categorical (factor) predictor, it has the advantage of allowing for a non-linear or threshold effect of increasing treatment level, and provides an unconstrained estimate of plasma ChE inhibition for each of the three dose levels. Model 2 treats dose level as a continuous quantity, and gives a single coefficient estimate that scales linearly to all doses; that is, the treatment effect estimates are constrained by the assumption that they all scale to the dose level by the coefficient estimated in the linear model. If the actual dose response is linear over the range of doses tested, the continuous dose model should give estimates that are both accurate and more precise than those

estimated by the categorical dose model. If the assumption of a linear dose response does not reflect reality, the categorical model would be expected to give more accurate slope estimates, albeit at the cost of lower precision.

**Model 1: dose as a categorical variable**

| | | | |
|---|---|---|---|
| $PChE_{ijt}$ | = | $\boldsymbol{\beta}_{0j}$ x **I**(dose = j) | *intercept = baseline mean* |
| | + | $\boldsymbol{\beta}_{1j}$ x **I**(dose = j) x $t_{trt}$ | *treatment effect* |
| | + | $\boldsymbol{\beta}_{2t}$ x **I**(day = t) | *date-specific effect* |
| | + | $\mathbf{b}_i$ | *subject random effect* |
| | + | $\varepsilon_{ijt}$ | *noise* |

$$b_i \sim \text{iid } N(0, \sigma^2_b), \quad \varepsilon_{ijt} \sim \text{iid } N(0, \sigma^2_{\varepsilon})$$

- $PChE_{ijt}$ is the plasma cholinesterase activity level in μmoles acetate/min/ml for subject i in dose group j on day t of the study.
- Dose j is categorical, identifying the dose groups 0, 14, 30, and 100μg/kg/day.
- Time $t_{trt}$ is a continuous measure of the number of days elapsed since the first treatment of the indicated dose group j. (Table 3.5 relates day of study t to treatment day $t_{trt}$)
- Day of study t is categorical, $t \in \{2,3,8,15,17,20,23,25,27,30,34,37,41,44,48,55\}$: This parameter allows for a different date-specific mean adjustment for day-to-day variation in lab results considered to be consistent across dose groups; the date-specific adjustment for baseline days (day 0) is included in the intercept $\boldsymbol{\beta}_{0j}$.

**Model 2: dose as a continuous variable:**

| | | | |
|---|---|---|---|
| $PChE_{ijt}$ | = | $\boldsymbol{\beta}_{0j}$ x **I**(dose = j) | *intercept = baseline mean* |
| | + | $\boldsymbol{\beta}_1$ x j x $t_{trt}$ | *treatment effect* |
| | + | $\boldsymbol{\beta}_{2t}$ x **I**(day = t) | *date-specific effect* |
| | + | $\mathbf{b}_i$ | *subject random effect* |
| | + | $\varepsilon_{ijt}$ | *noise* |

$$b_i \sim \text{iid } N(0, \sigma^2_b), \quad \varepsilon_{ijt} \sim \text{iid } N(0, \sigma^2_\varepsilon)$$

- $PChE_{ijt}$ is the plasma cholinesterase activity level in μmoles acetate/min/ml for subject i in dose group j on day t of the study.
- Dose j is a continuous quantity measured in μg/kg/day; dose-specific intercepts are defined only for $j \in \{0,14,30,100\}$ ; dose-specific treatment effects scale to all possible values of j.
- Time $t_{trt}$ is a continuous measure of the number of days elapsed since the first treatment of the indicated dose group j. (Table 3.5 relates day of study t to treatment day $t_{trt}$)
- Day of study t is categorical, $t \in \{2,3,8,15,17,20,23,25,27,30,34,37,41,44,48,55\}$: This parameter allows for a different date-specific mean adjustment for day-to-day variation in lab results considered to be consistent across dose groups; the date-specific adjustment for baseline days (day 0) is included in the intercept $\boldsymbol{\beta}_{0j}$.

## 2.5 Results of treatment analysis by LMM regression

Figure 2.3 overlays the smoothed mean results for all dose groups with the trend lines estimated by our two models. Each mean data point has been adjusted by subtraction of the day-specific mean effect computed in the model -- this permits valid comparison between the mean data and the modeled trends that have also been adjusted for the day-specific effect in the regression. The categorical model fits the data trend very closely (Figure 2.3a). *Slopes* estimated by the continuous model appear to reasonably approximate the rates of change in activity over time, but overall fit is clearly worse that that achieved by the model that allows both intercepts and slopes to vary without a linear constraint (Figure 2.3b). Table 2.3 compares the treatment effect slope estimates computed by the categorical- and continuous dose models. Slope estimates for a given dose differ only moderately between the two models, as demonstrated by the

considerable overlap between their 95% confidence intervals. The fact that constraining the slope estimates within a linear relationship (in Model 2) does not appreciably affect their values (relative to Model 1) suggests that the assumption of a linear relationship between daily dose and the rate of plasma ChE inhibition is reasonable.

When all measurements are included in a model that treats dose as a categorical predictor, the data show strong evidence that plasma ChE activity declines over time with daily treatment at the 30µg/kg and 100µg/kg levels ($p < 0.0001$). There is suggestion that plasma ChE activity also declines with daily doses of 14µg/kg, but the result is not statistically significant ($p = 0.065$). In contrast, the continuous dose model estimates a significant non-zero inhibitory effect at all three dose levels.

Figure 2.3 Categorical and continuous dose model fit to smoothed and lab-adjusted means

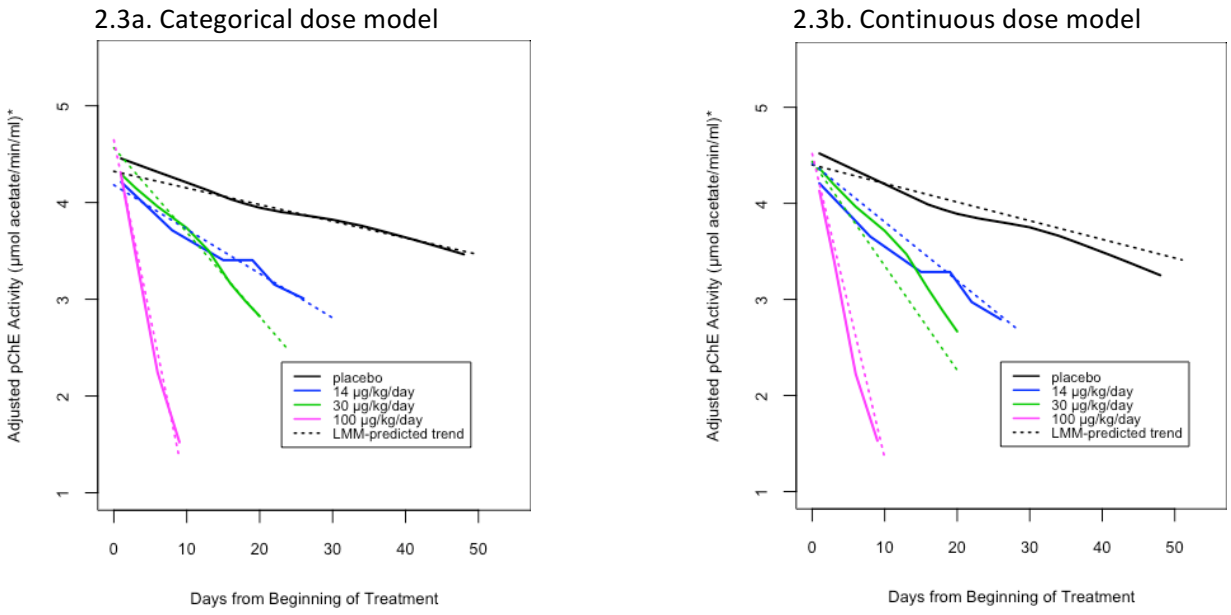


Table 2.4 Treatment effects estimated in µmoles acetate/min/ml/day by linear mixed effects regression

| dose (µg/kg/day) | treatment slopes* (95% lower, upper) | | | | | |
|---|---|---|---|---|---|---|
| | 14 | | 30 | | 100 | |
| categorical dose model | -0.029 | (-0.059, 0.002) | -0.070 | (-0.100, -0.040) | -0.349 | (-0.430, -0.270) |
| continuous dose model | -0.042 | (-0.089, -0.298) | -0.089 | (-0.093, -0.086) | -0.298 | (-0.304, -0.291) |

Both regression models indicate that plasma ChE activity declines more quickly with increasing dose level (Table 2.4). The categorical dose model estimates a trend of -0.017µmoles acetate/min/ml/day for the control group; this effect estimate does not differ significantly from 0, with p = 0.29 and 95% CI ( -0.049, 0.015). Among the subjects receiving chlorpyrifos, the estimated daily change of -0.070 µmoles acetate/min/ml/day for the 30µg/kg group is significantly more rapid than the corresponding estimate of -0.029µmoles acetate/min/ml/day for the 14µg/kg group (p=0.028). Similarly, the estimated daily change of -0.349µmoles acetate/min/ml/day for the 100µg/kg group is significantly more rapid than the estimate of -0.070µmoles acetate/min/ml/day for the 30µg/kg group (p<0.0001).

Treatment slope estimates from the continuous dose model are pre-specified to vary by dose, and 95% confidence intervals do not overlap between dose-specific effect estimates. In fact, this model's slope estimates have much narrower confidence intervals than those from the categorical model. This is true because fewer degrees of freedom are lost when specifying the single treatment slope coefficient in Model 2 vs. three coefficients in Model 1; therefore, standard errors are smaller in the continuous dose model. If a linear dose response can be assumed, the model estimates a control effect of -0.019µmoles acetate/min/ml/day; the estimate does not quite differ from 0, with p = 0.093 and 95% CI (-0.042, 0.003). Among treated subjects, the estimated rates of inhibition are 0.042µmoles acetate/min/ml/day at 14µg/kg/day, -0.089µmoles

acetate/min/ml/day at 30µg/kg/day, and -0.298µmoles acetate/min/ml/day at 100µg/kg/day, with $p<0.0001$ for the test of difference from zero in all three cases.

It is worth noting that both models appear to over-correct the control group measurements for daily confounding. As a consequence, controls trend estimates are exaggerated and slightly negative -- though neither estimate differs significantly from 0.

**2.6 Discussion**

Certain details of the study design are not documented in the available materials. Notably, it is not clear that subjects and investigators were blinded to treatment; given that treatment starting dates differed by dose group, it is likely that at least the investigators were not fully blinded. Also, because it is not specified that blood samples were collected at the same time on each collection day, the amount of time elapsed between each activity measurement and the most recent treatment could be inconsistent. For the analysis, it is assumed that the intervals were approximately uniform at 24 hours. Third, it is not explicitly stated that blood samples were immediately assayed for plasma ChE activity on the day of collection; hence the period by which lab results would be expected to vary is not known to be exactly one measurement day. Because the report takes care to mention that urine collected to measure DOWCO 179 and its metabolites was not immediately analyzed, we have assumed that immediate testing of samples was the norm.

The original analysis by repeated measures ANOVA correctly identified a significant reduction in plasma ChE activity with daily ingestion of chlorpyrifos at 100µg/kg. However, the

unexplained omission of one day's baseline data in the 30µg/kg dose group analysis led to an incorrect determination of no significant reduction in plasma ChE activity. With all baseline data included, there is a significant inhibitory effect of treatment at 30µg/kg/day (p=0.012).

Because starting dates and durations of treatment varied between the dose groups, the original data analysis did not have the technology to estimate differences in effect between treatment groups using repeated measures ANOVA; only comparisons between each treatment group and the controls were possible. This restriction eliminated the possibility to identify a graded association between dose level and treatment effect. Furthermore, the lack of balanced data required some valid measurements to be excluded from the analysis and allowed for judgment as to which data were included in the analyses. The removal of unbalanced data from an already small data set reduced the original study's power to identify a treatment effect, and the previously mentioned omission of an entire day's baseline data from the 30µg/kg dose group ANOVA led to a fundamentally different toxicological effect threshold being reported than was supported by the study data. The discrepancy would later impact regulatory policy, when the DOWCO 179 study's no-observed-effects-level (NOEL) became the basis for EPA's reference dose determination for chlorpyrifos.

## 3. Recovery analysis

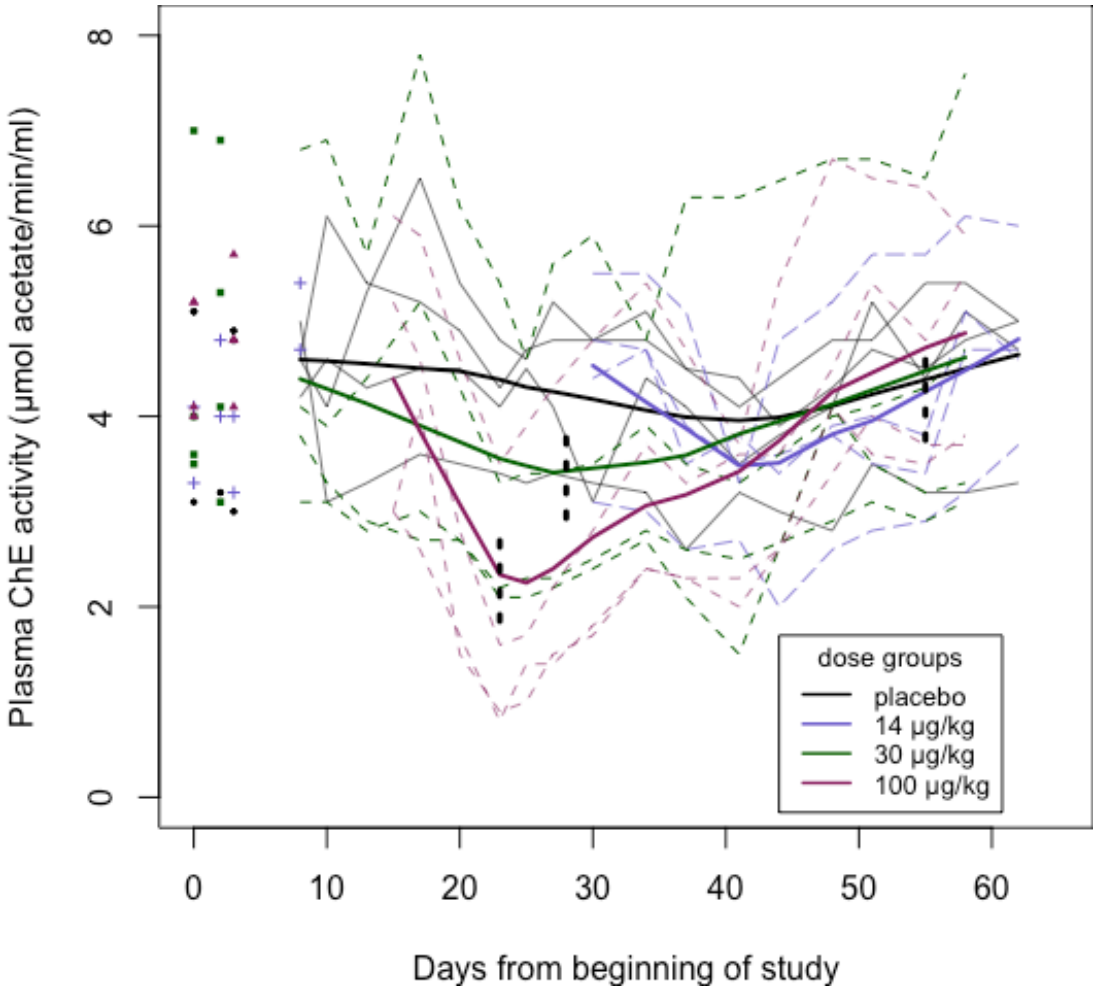
### 3.1 Original assessment of recovery

Following cessation of treatment, the DOWCO 179 safety evaluation continued to monitor all subjects until their cholinesterase activity returned to pre-treatment levels. In addition to recording semi-weekly measurements of RBC-ChE and plasma ChE activity, the researchers tracked weekly changes in subjects' serum chemistry, hematology, and urinalysis results. Though Coulston's report does not explicitly define the endpoint criterion, it appears that mean plasma ChE activity levels for each dose group, rather than individual subject activity levels, may have been used to make the determination that plasma ChE levels had returned to baseline levels. In fact, one subject in the 100μg/kg/day group did not fully regain his baseline activity by the study's completion. If statistical tests were used to decide whether activity values had returned to baseline levels, those tests were not described in the report.

Though the original analysis detected no reduction in plasma ChE activity for the group treated at 14μg/kg/day, the investigators did monitor those subjects for a one-week recovery period after treatments ended. The two higher dose groups, whose plasma ChE activity levels were measurably reduced by treatment, were followed for more than four weeks post-treatment. Coulston reports that recovery of plasma ChE activities began immediately after cessation of treatment, that no changes of toxicological significance were observed in RBC-ChE activity during recovery, and that all urine samples continued to test negative for DOWCO 179, its oxygen analog, and the decomposition product TCPy. Aside from noting their use of plasma ChE activity to establish an endpoint for the recovery period, the researchers did not report any quantitative analysis of plasma ChE activity levels measured during recovery. Figure 3.1 plots all

unadjusted individual and group mean plasma ChE measurements over the duration of the study, with vertical bars marking the treatment period endpoints defined in the Park report.

Figure 3.1 Unadjusted plasma ChE measurements: individual data and smoothed group means for both treatment and recovery periods. Vertical dashes intersecting mean lines indicate ends of treatment periods.




### 3.2 Selecting and indexing the data for recovery analysis

The objective of the recovery analysis was to estimate, for each dose group, the rate of recovery to normal plasma ChE activity following repeated daily ingestion of chlorpyrifos. Here, a single LMM regression model with splines fit descending trend lines to each dose group's treatment data and rising trend lines to recovery data for the 30µg/kg/day and 100µg/kg/day dose

groups. As in the preceding analysis of treatment effect, this analysis estimates the mean temporal trend in plasma ChE activity for each dose group, adjusting for subject-specific random effects and confounding by fixed effects from day-to-day variation in the assay results. An appropriate model should give dose-specific treatment effect estimates that are similar to those produced in the preceding treatment phase analysis, and the recovery effect estimates will be newly estimated rates of return to normal plasma ChE activity following inhibition.

Because the treatment/recovery model computes separate slope estimates for treatment and recovery effects, it is important to correctly classify each non-baseline measurement in the model as either treatment period or recovery period data. Given the small number of measurements available, misclassified data could have large impacts on the effect estimates. In addition, each measurement must be indexed by the correct number of days of treatment or recovery that the subject had undergone at the time of the measurement.

The indices for treatment period measurements were determined by the schedule of daily chlorpyrifos administrations. That is, our treatment period time indices are identical to those used in both Park's analysis and the preceding treatment analysis by LMM regression (see Section 2.4 for description.). The index associated with a treatment measurement indicates the number of daily chlorpyrifos treatments the subject had received one day prior to collection of the sample.

Timing indices for the recovery data were more challenging to define. The exact time at which toxicological recovery began could not be determined from the semi-weekly activity measurements, and Coulston did not cite evidence to support the reported observation that

recovery began immediately after treatment ended. Ultimately, we assigned indices for the recovery measurements using dataset annotations, best guesses based on outside research, and exploratory use of select measurement subsets in the regression model.

Our process for choosing indices began with evidence that maximal plasma ChE inhibition is typically seen minutes to hours after an acute chlorpyrifos exposure (Stefanidou, 2009). It seems reasonable, therefore, to conclude that a subject's recovery should have begun within the first 24 hours after the final treatment; however, because the time of blood collection was known only to the nearest day, we could not justify indexing measurements in units smaller than 1 day. Using Park's notes on the timing of treatments, we tentatively defined the time point one day after the last treatment to be both the last treatment day and recovery day 0. Later, exploratory use of the model to evaluate alternative starting points supported the initial choice (details are described in Section 3.3.2).

In our initial assessment of the data, we considered the recovery period endpoint to be the last day for which recovery measurements were provided. Our subsequent decisions to omit some of the later recovery measurements are explained below and in Section 3.3.1. Table 3.1 encodes all data as we assigned them within the final recovery model. In the table, baseline measurements are shaded in purple, treatment data are shaded in green, and recovery data are shaded in orange. The gray values are candidate recovery measurements that were ultimately omitted from the model. *A priori*, we judged that the control and 14µg/kg/day dose groups, with only two recovery days each, did not have sufficient data to reliably estimate recovery slopes; therefore, we excluded the recovery data (days 58 and 62) for those groups and only used the

model to estimate recovery trends for the 30µg/kg/day and 100µg/kg/day dose groups. Next, exploratory modeling described in Section 3.3.1 supported our decision to drop the last two measurements (days 55 and 58) for the 30µg/kg/day and 100µg/kg/day groups. After all treatment group measurements for days 55 and 62 had been omitted, we deemed the control measurements for those days to be superfluous and removed them from the model as well.

Table 3.1 Data classification in the final recovery model with splines

| | | control subjects | | | | 14µg subjects | | | | 30µg subjects | | | | 100µg subjects | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **date** | **day** | **2** | **4** | **9** | **16** | **7** | **8** | **14** | **15** | **10** | **11** | **12** | **13** | **1** | **3** | **5** | **6** |
| 29-Jun | 0 | 5.1 | 5.1 | 3.1 | 4.0 | 4.1 | 3.3 | | | 4.0 | 3.5 | 7.0 | 3.6 | 4.0 | 5.2 | 5.2 | 4.1 |
| 1-Jul | 2 | | | | 3.2 | | | 4.0 | 4.8 | 4.1 | 3.1 | 6.9 | 5.3 | | | | |
| 2-Jul | 3 | 4.8 | 4.9 | 3.0 | | 4.0 | 3.2 | | | | | | | 4.1 | 4.8 | 5.7 | 4.8 |
| 7-Jul | 8 | 4.6 | 4.6 | 5.0 | 4.2 | | | 5.4 | 4.7 | 4.1 | 3.1 | 6.8 | 3.8 | | | | |
| 9-Jul | 10 | 6.1 | 4.1 | 3.1 | 4.6 | | | | | 3.9 | 3.1 | 6.9 | 3.3 | | | | |
| 12-Jul | 13 | 5.4 | 5.3 | 3.3 | 4.3 | | | | | 4.4 | 2.8 | 5.7 | 2.9 | | | | |
| 14-Jul | 15 | | | | | | | | | | | | | 3.0 | 5.2 | 6.1 | 3.0 |
| 16-Jul | 17 | 5.2 | 6.5 | 3.6 | 4.5 | | | | | 5.2 | 3.0 | 7.8 | 2.7 | 4.1 | 4.6 | 5.9 | 2.6 |
| 19-Jul | 20 | 4.9 | 5.4 | 3.5 | 4.5 | | | | | 4.3 | 2.7 | 6.2 | 2.7 | 1.5 | 2.8 | 4.6 | 1.7 |
| 22-Jul | 23 | 4.3 | 4.8 | 3.4 | 4.1 | | | | | 3.3 | 2.2 | 5.4 | 2.1 | 0.9 | 1.6 | 3.5 | 0.8 |
| 24-Jul | 25 | 4.7 | 4.6 | 3.3 | 4.5 | | | | | 3.4 | 2.3 | 4.6 | 2.1 | 1.0 | 1.7 | 3.9 | 1.4 |
| 26-Jul | 27 | 4.8 | 5.2 | 3.4 | 4.1 | | | | | 3.4 | 2.3 | 5.6 | 2.2 | 1.5 | 2.2 | 4.3 | 1.4 |
| 29-Jul | 30 | 4.8 | 4.8 | 3.3 | 3.1 | 4.4 | 3.1 | 5.5 | 4.8 | 3.5 | 2.5 | 5.9 | 2.4 | 1.7 | 2.8 | 4.8 | 1.8 |
| 2-Aug | 34 | 4.8 | 5.1 | 3.2 | 4.4 | 4.7 | 3.0 | 5.5 | 4.7 | 3.9 | 2.8 | 4.8 | 2.7 | 2.4 | 3.7 | 5.4 | 2.4 |
| 5-Aug | 37 | 4.5 | 4.5 | 2.6 | 4.1 | 3.9 | 2.6 | 5.1 | 3.5 | 3.5 | 2.6 | 6.3 | 2.1 | 2.3 | 3.3 | 4.7 | 2.3 |
| 9-Aug | 41 | 4.4 | 4.1 | 3.2 | 3.5 | 3.5 | 2.7 | 3.3 | 3.8 | 3.3 | 2.5 | 6.3 | 1.5 | 2.3 | 3.6 | 3.7 | 2.0 |
| 12-Aug | 44 | 3.9 | 4.4 | 3.0 | 3.8 | 3.6 | 2.0 | 4.8 | 3.4 | | | | | 2.6 | 3.6 | 5.4 | 2.6 |
| 16-Aug | 48 | 4.3 | 4.8 | 2.8 | 4.1 | 3.8 | 2.6 | 5.2 | 3.9 | 4.0 | 2.9 | 6.7 | 4.1 | 4.1 | 4.5 | 6.7 | 3.4 |
| 19-Aug | 51 | 4.7 | 4.8 | 3.5 | 5.2 | 3.5 | 2.8 | 5.7 | 4.0 | 4.1 | 3.1 | 6.7 | 3.5 | 3.6 | 5.4 | 6.5 | 4.0 |
| 23-Aug | 55 | 4.5 | 5.4 | 3.2 | 4.4 | 3.4 | 2.9 | 5.7 | 3.8 | 4.3 | 2.9 | 6.5 | 3.2 | 3.5 | 4.8 | 6.4 | 3.7 |
| 26-Aug | 58 | 4.8 | 5.4 | 3.2 | 5.1 | 4.7 | 3.2 | 6.1 | 5.1 | 4.7 | 3.1 | 7.6 | 3.3 | 3.8 | 5.5 | 5.9 | 3.7 |
| 30-Aug | 62 | 5.0 | 5.0 | 3.3 | 4.7 | 4.7 | 3.7 | 6.0 | 4.6 | | | | | | | | |

baseline measurements
treatment measurements
recovery measurements used in final analysis
recovery measurements omitted from final analysis

### 3.3 Recovery analysis by linear mixed effects regression with splines

The treatment/recovery model (Model 3 below) is encoded using 3 timing indices: 2 expressions of continuous time ($t_{trt}$ and $t_{rec}$) in the day*dose expression, plus the categorical 'day

of study' index (t) for day-to-day variability. Two continuous time variables are required so that the model will compute distinct slope estimates for treatment and recovery. Coding of the treatment effect variable is identical to that used in the treatment model, except that the time index $t_{trt}$ takes on values larger than the index of the last treatment day; its coefficient is the estimated treatment slope. The recovery effect variable is coded as 0 for all days prior to recovery, so the model does not use it to predict the treatment trend. The variable becomes an index of recovery days for study dates beyond the knot marking the onset of recovery, and its coefficient, when added to the treatment slope coefficient, estimates the rate of recovery. Table 3.2 illustrates how the three indices of time used in the treatment/recovery model relate to the calendar date.

**Model 3: The treatment/recovery model**

$$
\begin{aligned}
PChE_{ijt} = \quad & \boldsymbol{\beta}_{0j} \times \mathbf{I}(\text{dose} = j) && \textit{intercept} \\
+ \quad & \boldsymbol{\beta}_{1j} \times \mathbf{I}(\text{dose} = j) \times t_{trt} && \textit{treatment effect} \\
+ \quad & \boldsymbol{\beta}_{2j} \times \mathbf{I}(\text{dose} = j \in \{30,100\}) \times t_{rec} && \textit{recovery effect} \\
+ \quad & \boldsymbol{\beta}_{3t} \times \mathbf{I}(\text{day} = t) && \textit{date-specific effect} \\
+ \quad & \mathbf{b}_i && \textit{subject random effect} \\
+ \quad & \varepsilon_{ijt} && \textit{noise}
\end{aligned}
$$

$$b_i \sim \text{iid } N(0, \sigma^2_b), \quad \varepsilon_{ijt} \sim \text{iid } N(0, \sigma^2_\varepsilon)$$

- $PChE_{ijt}$ is the plasma cholinesterase activity level in μmoles acetate/min/ml for subject i in dose group j on day t of the study.
- Dose j is categorical, identifying the dose groups 0, 14, 30, and 100μg/kg/day.
- Time $t_{trt}$ is a continuous measure of the number of days elapsed since the first treatment of the indicated dose group j.
- Time $t_{rec}$ is a continuous measure of the number of days elapsed since the last treatment day for the indicated dose group j.
- Day of study t is categorical, $t \in \{2,3,8,15,17,20,23,25,27,30,34,37,41,44,48,55\}$

Table 3.2 Time expressions in recovery model: $t_{trt}$ and $t_{rec}$ correspondence to date and day of study t

| day index | | control group | | 14µg group | | 30µg group | | 100µg group | |
|---|---|---|---|---|---|---|---|---|---|
| date | t | $t_{trt}$ | $t_{rec}$ | $t_{trt}$ | $t_{rec}$ | $t_{trt}$ | $t_{rec}$ | $t_{trt}$ | $t_{rec}$ |
| 29-Jun | 0 | 0 | | 0 | | 0 | | 0 | |
| 1-Jul | 2 | 0 | | 0 | | 0 | | | |
| 2-Jul | 3 | 0 | | 0 | | | | 0 | |
| 7-Jul | 8 | 1 | | 0 | | 1 | | | |
| 9-Jul | 10 | 3 | | | | 3 | | | |
| 12-Jul | 13 | 6 | | | | 6 | | | |
| 14-Jul | 15 | | | | | | | 1 | |
| 16-Jul | 17 | 10 | | | | 10 | | 3 | |
| 19-Jul | 20 | 13 | | | | 13 | | 6 | |
| 22-Jul | 23 | 16 | | | | 16 | | 9 | |
| 24-Jul | 25 | 18 | | | | 18 | | 11 | 1 |
| 26-Jul | 27 | 20 | | | | 20 | | 13 | 3 |
| 29-Jul | 30 | 23 | | 1 | | 23 | 1 | 16 | 6 |
| 2-Aug | 34 | 27 | | 5 | | 27 | 5 | 20 | 10 |
| 5-Aug | 37 | 30 | | 8 | | 30 | 8 | 23 | 13 |
| 9-Aug | 41 | 34 | | 12 | | 34 | 12 | 27 | 17 |
| 12-Aug | 44 | 37 | | 15 | | | | 30 | 20 |
| 16-Aug | 48 | 41 | | 19 | | 41 | 19 | 34 | 24 |
| 19-Aug | 51 | 44 | | 22 | | 44 | 22 | 37 | 27 |
| 23-Aug | 55 | 48 | | 26 | | | | | |
| 26-Aug | 58 | | | | | | | | |
| 30-Aug | 62 | | | | | | | | |

baseline day
treatment day
recovery day
unused recovery data

t day of study
$t_{trt}$ treatment day coding
$t_{rec}$ recovery day coding

### 3.3.1 Model sensitivity to recovery endpoint

Visual inspection of the mean data (Figures 3.2a and 3.3a) suggests that subjects in both the 30µg/kg/day and 100µg/kg/day groups might have regained their mean baseline plasma ChE activity before the last activity measurements were taken on August 26th. If so, fitting a straight line to data between the knot and the last data point could underestimate the recovery slope. Given sufficient data, a complete model of a plasma ChE inhibition and recovery could encode separate splines for treatment, recovery, and post-recovery baseline. In this case, however, the

small number of late-study measurements does not support an estimate of the post-recovery trend. As an alternative, we chose to run the model repeatedly on subsets of the data, sequentially deleting measurements at the end of the recovery period. We then used spline fit and the pattern of variation in the recovery estimate to optimize our data selection. The knot, or transition point between treatment and recovery splines, was provisionally set on the day following the last administration of treatment for each dose group. While this is not the optimal approach to determining the end of recovery, given the limited data, we believe it is the only feasible approach for the dataset.

For the 100µg group, sequential deletion of late recovery day measurements resulted in an increase in the slope of the modeled recovery line until it stabilized near 0.10 after the last two days' data were deleted (Table 3.3). Figure 3.1 permits visual comparison between the model fit using all data points and that omitting data for days 55 and 58. Mean data in the plot have been adjusted for day-to-day variation, and individual data have been corrected for both day-to-day variation and subject-specific mean variation. With the last two data points deleted, the model-predicted lines fit the data much more closely. Though also omitting day 51 data resulted in a further small increase in the recovery slope estimate, retaining that day seems prudent given the good fit seen in Figure 3.1b and the more precise estimate afforded by estimating the recovery slope from a larger data set. By similar reasoning, omitting only recovery days 58 and 55 for the 30µg/kg/day group seemed to optimize the balance between a good model fit to the data (Table 3.4, Figure 3.2) and the desire to retain as many data points as possible.

Table 3.3 Recovery slope dependence on chosen endpoint: 100µg dose group

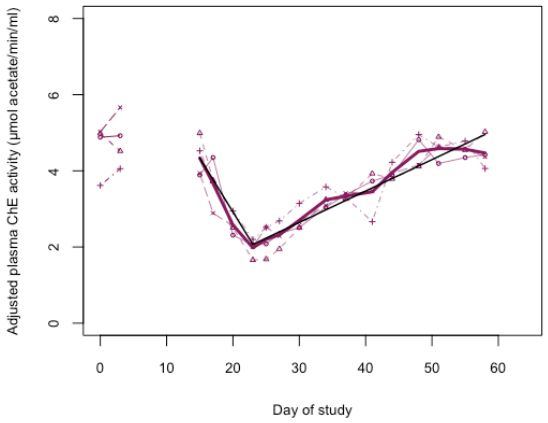
| Last recovery day | TRTMT slope estimate | 95% LWR | 95% UPR | RCVRY slope estimate | 95% LWR | 95% UPR |
|---|---|---|---|---|---|---|
| **58** | **-0.283** | -0.345 | -0.222 | **0.082** | -0.006 | 0.171 |
| **55** | **-0.298** | -0.359 | -0.237 | **0.093** | 0.004 | 0.182 |
| **51** | **-0.311** | -0.373 | -0.249 | **0.099** | 0.008 | 0.190 |
| **48** | **-0.315** | -0.378 | -0.252 | **0.102** | 0.007 | 0.196 |
| **44** | **-0.310** | -0.375 | -0.246 | **0.098** | -0.002 | 0.198 |
| **41** | **-0.310** | -0.376 | -0.244 | **0.098** | -0.008 | 0.204 |

Table 3.4 Recovery slope dependence on chosen endpoint: 30µg dose group

| last day | n | MSE | treatment slope | 95% lower | 95% upper | recovery slope | 95% lower | 95% upper |
|---|---|---|---|---|---|---|---|---|
| **58** | 260 | 0.1707 | **-0.065** | -0.094 | -0.036 | **0.031** | -0.004 | 0.067 |
| **55** | 256 | 0.1697 | **-0.067** | -0.097 | -0.038 | **0.037** | 0.000 | 0.075 |
| **51** | 252 | 0.1701 | **-0.072** | -0.102 | -0.042 | **0.046** | 0.007 | 0.084 |
| **48** | 248 | 0.1711 | **-0.076** | -0.106 | -0.045 | **0.056** | 0.014 | 0.097 |

Figure 3.2 Model sensitivity to choice of endpoint: 100µg dose group

3.2a. All 100µg recovery days included




3.2b. Last two recovery days omitted

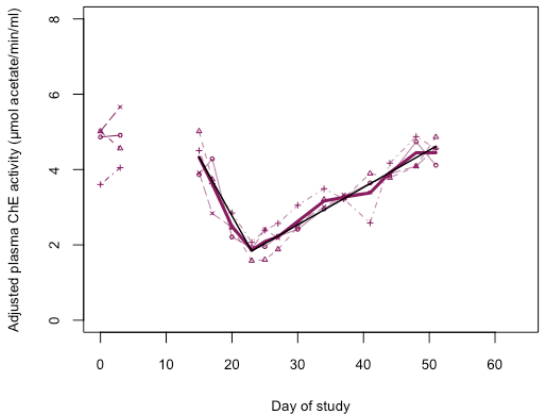
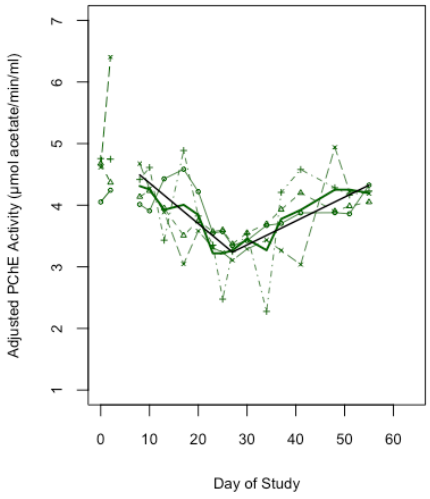
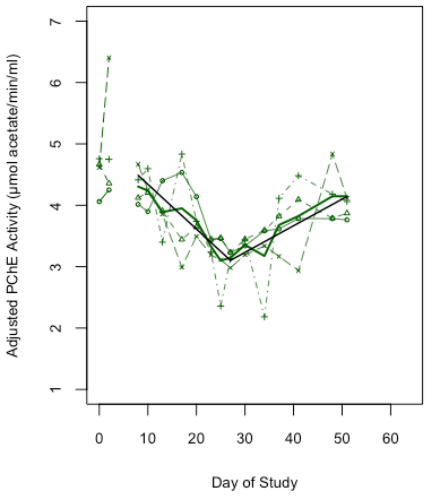


Figure 3.3 Model sensitivity to choice of endpoint: 30µg group

3.3a. All 30µg recovery days included




3.3b. Last two recovery days omitted




### 3.3.2 Model sensitivity to knot position

We also employed the treatment/recovery model to help to identify the most likely time point at which response to treatment transitions to recovery. Using our initial best guess, we started with the day following the last administration of DOWCO 17 and ran the model with the transition (knot) positioned before and after the starting point in increments of one day. The circled annotations in Appendix A are examples of the information used to locate the transition points. We then applied the following rules for selecting the preferred knot location:

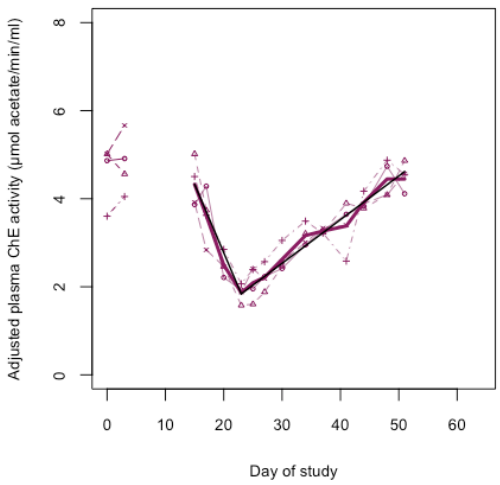
1. Choose the day that gives the best model fit by visual examination.
2. Verify that the chosen day gives a treatment effect estimate close to that estimated by the unconstrained treatment model.
3. Verify that the preferred model does not show obvious structure in residual plots.

This approach favored placement of the knot for both 30µg/kg/day and 100µg/kg/day dose groups on the day following the last treatment as documented by the original investigators (see Appendix A). For that knot position there is good model fit to the data (Figure 3.4) and the

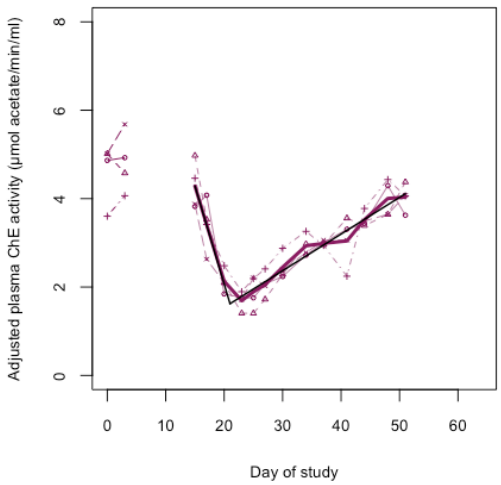
treatment slope estimates are very similar to those in the unconstrained model (Table 3.5). Residual plots did not have much diagnostic utility in this case, as none of the candidate models show obvious structure in their residuals.

Figure 3.4 Model fit with knot positioned at end of treatment (panels a and c) and 1 day before end of treatment (Panels b and d)

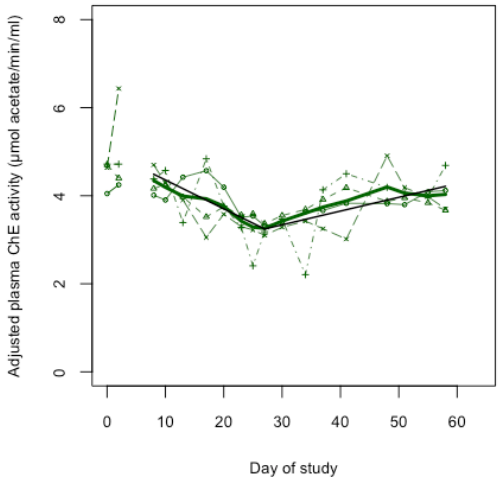
3.4a.




3.4b.




3.4c




3.4d

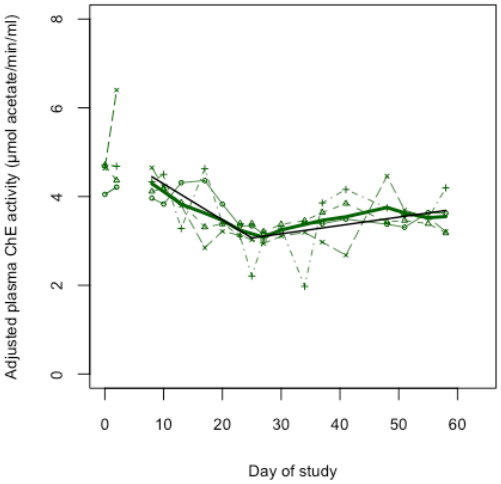


Table 3.5 Full model slope estimates for treatment and recovery

| | unconstrained model slope estimate | spline model slope estimate | 95% CI lower | 95% CI upper |
|---|---|---|---|---|
| **control trt** | -0.017 | -0.009 | -0.038 | 0.020 |
| **14µg trt** | -0.029 | -0.028 | -0.078 | 0.022 |
| **30µg trt** | -0.070 | -0.073 | -0.103 | -0.042 |
| **100µg trt** | -0.349 | -0.321 | -0.385 | -0.257 |
| **14µg rcvry** | - | - | - | - |
| **30µg rcvry** | - | 0.043 | 0.004 | 0.083 |
| **100µg rcvry** | - | 0.101 | 0.065 | 0.137 |

### 3.4 Results of recovery analysis

Raw output from the final treatment/recovery model can be found in Appendix D. Treatment slope estimates can be read directly from the output as the fdose:dft coefficients; as in the treatment model, these represent rates of change in plasma ChE activity (in µmoles acetate/min/ml/day) relative to the presumed actual control slope of zero. Recovery slopes are computed as the sum of the treatment slope estimate and recovery slope estimates for each dose group. For example, the recovery slope estimate for the 30µg/kg/day dose group is the sum of the fdose30:dft and recov30 coefficients. Standard errors for the sum are computed using the relationship $\text{var}(\beta_x, \beta_y) = \text{var}(\beta_x) + \text{var}(\beta_y) + 2\text{cov}(\beta_x, \beta_y)$, with input from the model's variance/covariance matrix.

According to the treatment/recovery model, treatment at both 30µg/kg/day and 100µg/kg/day clearly resulted in depressed plasma ChE activity, with depression occurring more rapidly in the 100µg/kg/day group. In units of µmol acetate/min/ml per day, slope estimates are -0.073 (95% CI -0.103, -0.042) for treatment at 30µg/kg/day, and -0.321 (95% CI -0.385, -0.257) for

treatment at 100µg/kg/day. Recovery slope estimates for the 30µg/kg/day and 100µg/kg/day groups are 0.043 (95% CI 0.004, 0.083) and 0.101 (95% CI 0.065 0.137), respectively, about one half and one third of the negative treatment slopes. That is, the data show that plasma ChE activity recovered more rapidly in the individuals whose enzyme activity had been more substantially reduced following chlorpyrifos ingestion. The difference between slope estimates is significant at $p = 0.018$.

Using these slope values, it is possible to approximate the amount of time required for each dose group's plasma ChE activity to return to baseline following treatment. For the 30µg/kg/day subjects, whose plasma ChE activity was depressed a model-estimated 25% by treatment, recovery would be expected to require 23 days (95% CI 12, 170). For the 100µg/kg/day subjects whose activity was depressed 62%, recovery would be expected to proceed more rapidly, with activity returning to baseline in 28 days (95% CI 21, 44).

**3.5 Discussion of recovery analysis results**

As seen in the treatment analysis of Section 2, the rate of plasma ChE inhibition during treatment is approximately proportional, at the tested levels, to the dose. The hypothesis of a fixed relationship between daily dose and *recovery* rate, however, does not seem plausible from a toxicological perspective. The biological half-life for ingested chlorpyrifos in humans is estimated at only 15.5 hours (Christensen, 2009), and the DOWCO researchers found no urinary traces of chlorpyrifos or its metabolites for any dose group, even among samples collected close to the time of greatest plasma ChE inhibition. Still, the recovery of enzyme activity continued for more than 3 weeks after the final treatments, in the likely absence of chlorpyrifos and its oxon.

Because recovery of plasma ChE activity proceeds much more slowly than elimination of the inhibitory chemicals, it seems likely that the rate-determining process in recovery involves reversal of entrenched enzyme inhibition or replacement of inactivated enzyme molecules. By this reasoning, the initial degree of plasma ChE inhibition could be expected to predict recovery rate better than daily dose, as it more directly quantifies the physiological demand that must be met during recovery.

In the DOWCO study, for subjects treated at 100µg/kg/day, mean plasma ChE activity had been reduced a model-estimated 62% below mean baseline by the last treatment day. Their recovery rate was estimated at 0.100 µmol acetate/min/ml/day, or 0.0016 µmol acetate/min/ml/day/% depression. By comparison, after 21 days of treatment at 30µg/kg/day, the lower dose group's mean plasma ChE activity had been reduced only an estimated 25% below mean baseline, and recovery was estimated at 0.047 µmol acetate/min/ml/day, or 0.0019 µmol acetate/min/ml/day/% depression. Because of high uncertainty in the slope estimates, calculations using them can support only speculative conclusions; however, the similarity between these ratios does not discredit the hypothesis that the recovery rate scales linearly to the degree of enzyme inhibition.

The observed difference in rates of recovery may be related to differential engagement of compensatory mechanisms between people with mild ChE inhibition and people whose ChE activity is more severely depressed. Examples of slow recovery mechanisms include dephosphorylation of the ChE binding site serine and up-regulation of ChE transcription, while faster recovery is facilitated by release of ChE molecules in the liver (Dr. Lucio Costa, personal

communication, August14, 2017). A consequence of the difference in recovery rates between the groups is that the subjects whose plasma ChE activity was less severely compromised had depressed activity for almost as long, though at a lower mean degree of inhibition, as the subjects who had been more acutely affected by treatment. That is, if *duration* determines the severity of the toxicological effect, a moderate dose may be as damaging as a large dose.

## 4. DOWCO 179 study's contribution to chlorpyrifos regulation in the United States

### 4.1 Chlorpyrifos regulation 1965-1996: FIFRA

Since its enactment in 1947, FIFRA has been the primary federal statute governing pesticide use in the US. FIFRA initially granted control of the registration process to the US Department of Agriculture (USDA), authorizing the Department to register a pesticide for interstate commerce if it could be proven effective for specified uses and unlikely to cause unreasonable harm to the environment. Congress transferred administration of FIFRA to EPA upon its creation in 1970. Following passage of the Environmental Pesticide Control Act in 1972, EPA's pesticide registration protocol shifted to place greater emphasis on minimizing risks to human health and less on a chemical's efficacy (Oleskey, 2004).

#### 4.1.1 Reference dose definition and calculation

In making the safety determination required for FIFRA registration, EPA estimates a typical human exposure for the candidate pesticide or use and compares it to a reference dose (RfD). The RfD is judged to be a level of daily oral exposure to the chemical that will cause no adverse health effects over a lifetime. It is also common for EPA to report an acute RfD, or level that should cause no adverse health effects after a single short-term exposure. Because pesticide exposures may occur via a combination of ingestion, inhalation, and dermal absorption, RfDs are usually expressed in readily converted units of pesticide mass per unit body weight per unit time.

To compute the RfD, EPA must first determine a pesticide's critical effect. This may be either the health effect of greatest concern, or that observed at the lowest level of exposure. To translate the critical effect into an action level, a toxicological endpoint must be identified. The

toxicological endpoint is a specific, measurable outcome associated with the critical effect. For example, in the case of an organophosphate pesticide, the critical effect might be respiratory arrest due to cholinergic response, and the toxicological endpoint could be a degree of enzyme inhibition previously shown to coincide with the onset of cholinergic symptoms. The point of departure (PoD) for a reference dose calculation is the highest level of exposure shown to have no toxicological endpoint effect in a reference study; this value is referred to as the No-Observed-Effects-Level (NOEL) or No-Observed-Adverse-Effects-Level (NOAEL). Pesticides that are known to be carcinogenic are ineligible for registration under FIFRA, as it is not possible to determine a level of exposure that poses no risk of long-term adverse effects (US EPA, 2002).

Because the RfD represents a level of exposure deemed safe for all members of the population, it is lower than the empirically determined NOEL used as the RfD's point of departure. Typically, to derive a chemical's RfD using human toxicology data, the NOEL is reduced 10-fold by an intra-species uncertainty factor (UF) to allow for variation in response between individuals. It may be further reduced by a safety factor (SF), typically set at 10, if especially sensitive groups could be exposed to the chemical; in that case, the resulting RfD value is 1/100 of the NOEL. If the starting point NOEL comes from an animal toxicology study, it is reduced 10-fold by an inter-species UF, 10-fold by the intra-species UF, and possibly another 10-fold if a margin of safety (SF) is required to protect sensitive groups; this yields an RfD that is 1/1000 of the NOEL. When available data support alternative uncertainty and safety factor levels, they may replace the default value of 10 (US EPA, 2002).

### 4.1.2 Reference dose for chlorpyrifos

Chlorpyrifos was first registered under FIFRA on July 1, 1965 for the control of foliage- and soil-borne insect pests on a variety of crops used in food commodities and animal feed. Its range of registered applications gradually expanded to include pet treatments, golf course pest suppression, non-structural wood treatments, industrial site pest control, and some residential use by consumers. In 1988, after chlordane's registration was revoked due to its potential carcinogenicity, chlorpyrifos became a first-line structural treatment for residential termite infestations (US EPA, 2016b).

EPA Health Effects Division's RfD/Peer Review Committee first established a dietary RfD for chlorpyrifos on February 21, 1986, using acute cholinergic response as its critical effect, plasma-ChE inhibition as the toxicological endpoint, and the repeated-dose NOEL from the 1972 Coulston study as its point of departure. The chosen NOEL value of 30µg/kg/day relied on Coulston's determination of no plasma-ChE inhibition after repeated treatments at that level, when significant plasma-ChE inhibition had been observed after 6 and 9 days of treatment at the next higher level of 100µg/kg/day. That presumed NOEL was then reduced by an inter-species UC of 10 to give the documented RfD of 3µg/kg/day. The value was affirmed a second time by EPA's RfD Committee on March 4, 1988, and remained unchanged following the Committee's assessment for FIFRA reregistration on September 8, 1993 (US EPA, 1998a).

## 4.2 Regulatory changes with 1996 enactment of FQPA

On August 3, 1996, with the unanimous support of the US Congress, President Clinton signed the Food Quality Protection Act (FQPA), a set of directives intended to improve health

standards applicable to food pesticide residues. FQPA's amendments to FIFRA fundamentally changed EPA's approach to pesticide regulation and explicitly elevated the protection of human health to top priority. As its main precept, FQPA instructed EPA to demonstrate in its setting of tolerances that a pesticide can be used with "reasonable certainty of no harm," a more stringent requirement than FIFRA's charge to show that using a pesticide according to specifications will "not generally cause unreasonable adverse effects on the environment." For risk assessments, FQPA required the Agency to estimate aggregate exposure from multiple sources and routes, and to consider cumulative risk for agents with common mechanisms of toxicity. Prior to FQPA, risk assessments had assessed risk from pesticide exposure on an individual, case-by-case basis. Additional FQPA provisions directed EPA to reassess, within 10 years of the Act's signing, all 9721 pesticide tolerances already in place; to schedule future registration reviews according to a 15-year cycle; and to improve methods for assessing pesticide use, food consumption patterns, and pesticide residue levels. Significantly, FQPA also stipulated that EPA's revised tolerance calculations must apply an additional safety factor (SF) between 1 and 10 to account for greater susceptibility to pesticides among infants and children (US EPA, 1998a; Oleskey, 2004).

### 4.2.1 Dow's arguments for higher NOEL and reduced margins of safety

The child-protective safety factor, to be set at 10-fold unless data supported use of a different value, would have reduced most existing pesticide RfDs to 1/10 of the levels computed prior to FQPA's enactment. Some registrants, hoping to avert or minimize a reduction in tolerance for their products, presented arguments for EPA to use alternative NOELs or reduced uncertainty and safety factors in computing RfDs. In the case of chlorpyrifos, Dow submitted a written petition to the EPA on April 25, 1995, well in advance of FQPA being signed into law, disputing

the Agency's use of plasma ChE inhibition as a toxicological endpoint. The petition proposed instead that RfD determinations use acute and chronic NOELs for RBC-ChE inhibition in humans (US EPA, 1998a).

Dow's 1995 letter asserted that plasma ChE has no known biological function in humans and is unrelated to adverse cholinergic health effects. Further, the petition contended that human and animal studies had demonstrated that chlorpyrifos inhibits plasma-ChE at much lower doses than those required to inhibit AChE in the nervous system, where cholinergic effects originate. Citing evidence that RBC-ChE correlates closely with AChE levels in brain and muscle tissue in animal models, Dow proposed that NOELs for RBC-ChE inhibition should be used as the starting values in EPA's RfD calculations for chlorpyrifos. For a single exposure, Dow suggested an NOEL of 500µg/kg/day, the one-time dose administered in a 1982 study of six adult male volunteers who showed no subsequent change in RBC-ChE activity (US EPA, 1998a; Nolan, 1984). For repeated exposures, Dow proposed an NOEL of 100µg/kg/day, the highest dose administered in the 1972 Coulston study that saw no RBC-ChE inhibition at any dose level. As an alternative, the letter suggested the repeated-dose NOEL of 1000µg/kg/day measured in a study of brain AChE inhibition in rats.

The Hazard Identification Assessment Review Committee (HIARC) of EPA's Health Effects Division (HED) met on May 25, 1995 and November 23, 1995 to address Dow's petition, ultimately concluding that continued use of plasma-ChE inhibition as the endpoint for chlorpyrifos regulation was supported by the weight of evidence. While the Committee agreed that plasma-ChE has no known role in cholinergic transmission, it disputed the contention that its

activity is a poor predictor of ChE activity in brain and muscle tissues. To support its position, the Committee cited literature showing that plasma-ChE inhibition in organophosphate-treated rats correlates well with AChE inhibition in the brain, while RBC-ChE inhibition does not. HIARC clarified its objection to proposed use of the single-dose NOEL from Nolan's 1982 study, noting that subjects' plasma-ChE activity was inhibited 84% to 86% on the first day of the study, and RBC-ChE was inhibited 11% to 52% on the fourth day of the study -- even though no further pesticide had been administered – and it could therefore not be stated that the single treatment had had no effect on either RBC-ChE or plasma-ChE activity (US EPA, 1998a). As evidence that plasma-ChE inhibition in the Coulston study may have been associated with systemic cholinergic effects, the Committee noted that blurred vision, one of the symptoms shown by a subject in the highest dose group, is typical of the cholinergic response but not usually associated with cold or influenza. Regarding the suggested use of an NOEL for AChE in rat brain, HIARC indicated that EPA prefers to use human data when possible, to minimize uncertainty from inter-species extrapolation.

On May 9, 1997, Dow followed up its earlier petition by providing HIARC with a document entitled *Proposed Reference Dose (RfD) for Acute and Chronic Exposure to Chlorpyrifos Based on the Criteria Described by the Acute Cholinesterase Risk Assessment Task Force and the Available Animal and Human Data*. The report reviewed a variety of toxicology studies that Dow cited in support of its suggested NOELs and uncertainty factors for RfD determinations. HIARC considered the report and accompanying reference studies at its December 11, 1997 and January 28, 1998 meetings to reassess allowable dietary and residential/occupational exposures after reviewing toxicological endpoints, uncertainty and safety factors, and evidence for

increased sensitivity in infants and children. At that time, the Committee did not feel that the studies successfully discredited EPA's previous endpoint selection or the application of a child-protective safety factor. In the EPA memorandum dated February 2, 1998, HIARC summarized its conclusions from the meetings, reaffirming its recommendation to retain NOELs established in the Coulston study as the basis for chlorpyrifos tolerance calculations. In addition, using only hazard assessment as its rationale, the Committee supported use of the full 10-fold safety factor to protect infants and children. Its principal arguments stated that chlorpyrifos is a known neurotoxicant associated with organophosphate-induced delayed neuropathy (OPIDN) in humans and animals, that increased susceptibility in offspring to the effects of chlorpyrifos had been demonstrated in both human and animal studies, and that several previously approved uses of the chemical had potential to expose infants and children. The February 2, 1998 memorandum recommended acute and chronic dietary RfDs of 1µg/kg/day and 0.3µg/kg/day, respectively, or 1/10 of the previously established values. In that document, HIARC also proposed short-, intermediate-, and long-term dermal and inhalation exposure limits that were derived from the Coulston study's oral dose NOELs (US EPA, 1998b).

At a subsequent meeting on October 29, 1998, HIARC reviewed five additional studies of chlorpyrifos toxicity in adult, fetal, and neonatal rats, some of which suggested that sensitivity to chlorpyrifos did not vary appreciably between rats of differing life stages (US EPA, 1998b). The Committee members agreed that, though the study results were not fully consistent, elevated sensitivity to chlorpyrifos among young rats could not be discounted. However, HIARC also acknowledged that the dose levels used in the studies were unrealistically high, and that the routes of administration were not typical of human exposure. The HIARC conclusions published

on December 7, 1998 were less emphatic in their support of the 10-fold child-protective safety factor than the February 2, 1998 statement, with 4 members voting to retain the full value and 7 members voting to reduce it. The FQPA Safety Factor Committee was expected to decide on a final value during its upcoming risk characterization proceeding.

### 4.2.2 Use restrictions under FQPA: 1997, 2000 and 2002

Between 1987 and 1998, nearly half of the 21 to 24 million pounds of chlorpyrifos applied annually in the US went to non-agricultural uses, including treatment of school facilities, residences, parks, golf courses, industrial sites, ships, and trains. In January, 1997, the technical registrants voluntarily agreed with EPA to reduce indoor chlorpyrifos exposures to children and other sensitive groups by restricting specific non-agricultural uses. Product types canceled at that time included indoor broadcast treatments and foggers, direct pet treatments, and paint additives (US EPA, 2006).

As an intermediate step in the process of reregistration specified by FQPA, EPA released a preliminary assessment of human health risks associated with registered chlorpyrifos uses on June 8, 2000 (US EPA, 2002). Because the assessment of cumulative risk from all organophosphates had not yet been completed, the document did not include a reregistration eligibility decision; however, it did include interim RfDs addressing FQPA-mandated updates to safety factors and risk assessment protocols, as well as emerging concern over the use of data from human experimentation in regulatory processes. Significantly, the announcement indicated concern that several registered chlorpyifos uses posed an unacceptably high health risk to children. At that time, in cooperation with Dow Agrosciences, EPA initiated measures to

immediately reduce those child health risks associated with dietary intake and certain residential and non-residential applications. Specifically, all use of chlorpyrifos on tomatoes was discontinued, application to apples was restricted to pre-bloom and dormant trees, and use on grapes was limited to dormant vines. In addition, all products marketed for home use, except ant and roach bait products in child-resistant packaging, were scheduled to be eliminated by December 31, 2001. It was agreed that application of chlorpyrifos as a termiticide would be phased out by December 31, 2005 unless credible data could demonstrate that exposures to residents were acceptably low (US EPA, 2002).

In addition to preliminary assessments of health risk to humans, EPA's June, 2000 report presented results from a screening-level ecological risk assessment for chlorpyrifos. Risk quotients indicated that nearly all outdoor uses posed acute and reproductive risks to small mammals, birds, fish, and aquatic invertebrates. In 2002, to mitigate the ecological effects of agricultural and non-agricultural outdoor uses of chlorpyrifos, EPA reduced the allowable application rates and number of applications per season for alfalfa, citrus crops, corn, cotton, peanuts, soybeans, sorghum, sugars beets, sunflowers and tree nuts. The reduction also addressed concern over risk to pesticide handlers, forming part of a three-tiered strategy to lower occupational exposures via engineering controls, use of personal protective equipment, and reductions in pesticide concentration and application frequency.

### 4.3 Moratoria on use of data from human dosing studies, 1999 and 2000-2004

Following the passage of FQPA in 1996, the added safety factor for infants and children effectively reduced allowable exposures for many pesticides to one-tenth of their previous levels.

Faced with higher safety standards, some pesticide manufacturers conducted new small-scale toxicology studies on humans to see whether evidence might support elimination of the interspecies uncertainty factor. If human subjects could be shown to have sensitivity similar to that measured in test animals, it would be advantageous for a company to have the allowable exposure level for its product computed as the NOEL from a human toxicology study adjusted only by an intra-species uncertainty factor and the applicable safety factor (Oleskey, 2004).

The Nuremberg Code, established shortly after the end of World War II, established the ethical guidelines that applied to the DOWCO 179 research in 1971. The Code emphasized the legal doctrine of informed, voluntary consent, as well as the principles of beneficence and non-maleficence, by which investigators must justify a study as being likely to produce benefit for society or the subjects, and should take measures to minimize any harm to participants that is not balanced by potential benefit (Resnik, 2005). In 1991, EPA and 15 other federal agencies adopted the Common Rule as their ethical framework for research on human subjects. The Common Rule, based on resolutions in the 1979 Belmont Report, augmented the Nuremberg Code by requiring stricter oversight and documentation of compliance by research institutions. It also added protections for specific vulnerable groups of potential subjects, including pregnant women, fetuses, children, and prisoners (US EPA, 2008).

On July 1, 1998, the Environmental Working Group (EWG) published a report entitled *The English Patients: Human Experiments and Pesticide Policy*, an overview of several recent studies by pesticide companies alleged to have intentionally exposed volunteers to toxins in order to determine human NOELs for their products (EWG, 1998). One example in the report is

a 1998 study by Dow AgroSciences said to have compensated 60 Nebraska college students to swallow tablets containing high doses of chlorpyrifos (Kisicki, 1999). The EWG report noted that, while EPA-conducted and EPA-sponsored research on human subjects must satisfy the Common Rule, third-party research referenced in EPA's regulatory determinations is not required to do so. EWG recommended that EPA conduct a review of past and current human experimentation in the context of environmental policy making, impose an immediate moratorium on third-party human experimentation until an ethics review had been completed, and adopt policy guidelines and procedures for pesticide testing before relaxing the moratorium on the use of human experimentation to support pesticide registration. The report noted that many of the referenced studies included fewer than 15 subjects, and were therefore unsuitable to characterize the variability in response across the entire population. Further, EWG asserted, because all of the studies tested pesticide responses in adults, they could not be justified as attempts to satisfy new requirements under FQPA, whose recent procedural changes addressed concern over potential elevated toxicity in children (EWG, 1998).

In response to public and professional concern over ethical issues raised in the EWG report, EPA convened a joint meeting between its Scientific Advisory Board (SAB) and Scientific Advisory Panel (SAP) on December 10th and 11th, 1998. Goals of the meeting were to formulate ethical guidelines for pesticide testing in humans and prepare a report of the resolutions. Though the participants agreed on broad principles, some members refused to sign the draft guidelines. After a year-long stalemate, the SAB and SAP re-convened to continue their deliberations on Nov 30, 1999, and released a final report on September 11, 2000. The report emphasized that acceptable studies must be of sufficient size to give statistically meaningful

results, that the risks and benefits from the work should be distributed throughout the population, and that human dosing studies are generally not appropriate for determining NOELs if the intent is to eliminate the interspecies uncertainty factor. Two participants released a minority report contending that the September 11, 2000 document understated the risks to human subjects from intentional dosing, and that no limited study could provide information about safe levels of pesticide exposure, especially for children (Oleskey, 2004; US EPA, 2008).

Between 1999 and 2003, EPA's policy regarding acceptance of third-party human pesticide studies reversed multiple times. In late 1999, the Clinton Administration directed EPA to stop accepting data from pesticide industry studies conducted on humans; this order superseded the pending SAB/SAP report. The Bush Administration overturned the 1999 ban in November, 2001, again approving EPA's use of data from human testing. After reviewing three recent industry-sponsored human dosing studies, including the 1999 chlorpyrifos study by Kisicki, EPA announced another moratorium on such research effective December 14, 2001. On Feb 5, 2002, with the moratorium still in effect, EPA signed a contract with the National Research Council (NRC) of the National Academy of Sciences (NAS), agreeing to appoint an expert committee to examine the controversial ethical issues surrounding pesticide experimentation in humans. In June 2003, the US Court of Appeals of the District of Columbia reinstated EPA's right to consider third-party results of human pesticide testing on a case-by-case basis, stipulating that continued tolerance of the practice would depend on public comment and the upcoming announcement from NRC. On February 4, 2004, NRC concluded that EPA should be allowed to accept data from studies that expose volunteers to pesticides as long as ethical standards were met. Further, NRC recommended that EPA adopt the Common Rule for third party research data

and appoint a committee to review proposed third party research and previously collected data (US EPA, 2008).

As required under FQPA, EPA released its updated registration eligibility decision (RED) for chlorpyrifos on July 31, 2006, approving all previously registered uses except those eliminated in 1997, 2000, and 2005, when termiticide treatments were phased out. Release of the 2006 RED for chlorpyrifos marked the formal adoption of acute and dietary RfDs based on NOAELs from animal studies, and the first time that finalized RfD determinations made no reference to the 1972 Coulston study. The RfDs adopted in 2006 were the same interim values suggested in June, 2000 as benchmarks for exposure estimates in support of the 2006 reregistration. The new acute dietary RfD of 5µg/kg/day was computed using a UF of 100 and the NOAEL of 500µg/kg/day observed in an acute blood time course study on rats. For children and women aged 13 to 50, the RfD was reduced by a safety factor (SF) of 10, to 0.5µg/kg/day. The NOAEL determination considered two endpoints: significant plasma ChE inhibition and significant RBC-ChE inhibition 4 hours after exposure. The new chronic dietary RfD came from the NOAEL of 30µg/kg/day determined by a weight-of-evidence approach, using 4 long-term studies in dogs and rats, and one short-term study of developmental neurotoxicity in rats. The RfD calculation used a total UF of 100 for the general population to give 0.3µg/kg/day, and an additional SF of 10 for children and women aged 13 to 50, yielding the RfD of 0.03µg/kg/day for those groups (US EPA, 2006). Table 4.1 summarizes the RfD levels and bases from the initial setting of a tolerance to the most recent reregistration eligibility decision in 2006.

Table 4.1 RfD values for chlorpyrifos since 1986

| Date | RfD | Population Adjusted Dose | Toxicological Endpoint | Reference Studies | Total Uncertainty Factor | FQPA Safety Factor |
|---|---|---|---|---|---|---|
| **21-Feb-86** | **3µg/kg/day** | - | plasma ChE inhibition | human repeated dosing study (Coulston, 1972) | 10 | 1 |
| **2-Feb-98** | **0.3µg/kg/day** | - | plasma ChE inhibition | human repeated dosing study (Coulston, 1972) | 10 | 10 |
| **31-Jul-06 (interim adoption 8-Jun-00)** | **0.3µg/kg/day** | - | plasma and RBC-Che inhibition at LOAEL of 220 to 300µg/kg/day | weight of evidence: 2 year dog, 90 day dog, 2 year rat, 90 day rat, developmental neurotoxicity in rat | 1000 | 1 |
| | - | **0.03µg/kg/day for children and females 13-50** | | | 1000 | 10 |

1986 and 1998 RfDs reported in US EPA 1998a. Chlorpyrifos - FQPA Requirement - Report of the Hazard Identification Assessment Review Committee. Office of Pesticide Programs. Washington, DC, February 2.

2006 (and interim 2000) RfDs reported in US EPA 2006. Reregistration Eligibility Decision for Chlorpyrifos. Office of Pesticide Programs. Washington, DC, July 31.

### 4.3.1 HSRB review of chlorpyrifos dosing studies

On February 21, 2006, EPA chartered an independent Human Studies Review Board (HSRB) to "provide advice, information, and recommendations on issues related to scientific and ethical aspects of human studies research." The HSRB would be charged specifically with providing advice on ethical aspects of proposed research, completed studies on human subjects, and EPA's policy for protecting human subjects. EPA announced its updated rules for protection of subjects in human research on April 7, 2006 (Resnik, 2007).

For data collected prior to April 7, 2006, study protocols would be subjected to HSRB review if the research intended to identify or measure a toxic effect. Data from certain types of previously completed pesticide research were declared inadmissible for EPA use: studies on pregnant women, nursing women, or children; studies whose conduct was fundamentally unethical; and studies whose conduct was deficient relative to the ethical standards prevailing at the time. Under the new rule, data from existing pesticide dosing studies, including the Coulston study of 1972, could be used to inform regulatory actions if the protocols received HSRB approval (Resnik, 2007).

By December 17, 2008, EPA's SAP had reviewed just three human studies of chlorpyrifos: Nolan (1982), Griffin (1999), and Kisicki (1999). The SAP determined that those studies were inappropriate for setting PoDs or UFs for the following reasons (US EPA 2009):

a. Laboratory studies in animals had suggested that both cholinergic and non-cholinergic mechanisms may contribute to developmental effects associated with exposure to chlorpyrifos. Epidemiologic studies in humans supported findings from the animal studies. The Nolan (1982), Griffin (1999), and Kisicki (1999) studies did not consider toxic endpoints other than ChE inhibition, and therefore were not relevant to protecting pregnant women and children.
b. The HSRB had previously criticized the study design used in Nolan (1982) and Griffin (1999) for considering only a single dose group treated via a single route of exposure.

c. The HSRB questioned the efficacy of the study design in Griffin (1999) and Kisicki (1999), as the investigators had not detected an effect (LOAEL) at any dose level.

Because no technical registrant for chlopyrifos submitted the 1972 Coulston study for SAP/HSRB review, it has become unavailable for consideration in EPA's regulatory decisions.

### 4.3.2 Previous assessment of the DOWCO 179 safety evaluation study

Though it has not appeared in recent regulatory determinations, the Coulston study did undergo a data evaluation by the Toxicology and Hazard Assessment Group at Oak Ridge National Laboratory in 2009 (Oak Ridge, 2009). The resulting document, dated May 26, updated the study's conclusions and identified the strengths and limitations of its design. Among the study's noted strengths were its capacity to provide dose-response estimates -- unlike the studies by Nolan, Griffin, and Kisicki, Coulston's work included multiple dose levels, a concurrent negative control, and blood sampling at successive time points. Further, Coulston measured both RBC-ChE and plasma ChE activity for all subjects, and tested for the presence of chlorpyrifos and its metabolites in urine. Last, the review noted that the study followed its subjects for the full duration of both treatment and recovery periods. Limitations cited in the review included many also noted in this thesis. Specifically, it was observed that each dose was given for a different amount of time without a clear rationale, ChE activities were not measured every day, and the amount of time elapsed between treatment and sample collection was not specified. In addition, the review faulted the study for omitting technical details on the analytical limits of detection and

nature of the test material, and it disapproved of experimentation at a correctional facility. Finally, the review disputed the original investigators' NOAEL determination of 30µg/kg/day for plasma ChE. Based on inhibition to 32% of baseline activity by day 16 of treatment at 30µg/kg/day, and continued depression of approximately 30% through the end of treatment on day 20, the reviewers considered 30µg/kg/day to be the LOAEL, and the next lower dose of 14µg/kg/day to be the NOAEL. That is, by applying a threshold criterion for plasma ChE depression, the Oak Ridge group arrived at the same NOAEL value suggested by our repeated measures ANOVA and categorical-dose LMM regression analyses.

### 4.4 PANNA/NRDC petition and risk assessment updates, 2007 to present

In 2007 the Pesticide Action Network of North America (PANNA) and the Natural Resources Defense Council (NRDC) petitioned EPA to revoke all tolerances for chlorpyrifos and cancel all uses registered under FIFRA (US EPA, 2017). In support of their position, the petitioners cited evidence from three prospective birth cohorts suggesting that low-level exposures *in utero* could adversely impact fetal development and later cognitive development in children (Whyatt, 2004; Rauh, 2006). Between 2008 and 2012, the SAP convened five times, but was unable to deliver a recommendation to either approve or deny the proposed revocation. In 2014, the petitioners brought suit in the US 9th Circuit Court of Appeals, seeking to compel EPA to deliver either a denial of their request or a tolerance revocation. On June 10, 2015, the 9th Circuit Court ordered EPA to inform the Court of its planned response to the petition, whereupon EPA announced its intention to release a proposed ruling on tolerance revocation by April 15, 2016. The Court rejected EPA's timeline, and insisted instead that the Agency produce a proposed or final ruling on

the petition by October 31, 2015. On October 30, 2015, EPA announced its preliminary intent to revoke tolerances as requested. Ultimately, on August 12, 2016, PANNA and NRDC secured orders through the 9th District Court requiring EPA to expedite its risk assessment and issue a final ruling on the petition no later than March 31, 2017. On March 29, 2017, newly appointed EPA Administrator Scott Pruitt formally denied the request for revocation of tolerances (US EPA, 2017). Still more recently, on June 5, 2017, PANNA and NRDC countered with a formal objection to the denial, requesting a reversal of the ruling within 60 days (NRDC, 2017). Pruitt has announced his intent to delay further response to the petition until the reregistration deadline in 2022.

Meanwhile, in 2009, EPA initiated risk assessments for the 2022 registration deadline. Chlorpyrifos and other organophosphates were moved forward in the schedule because of ongoing concern that neurodevelopmental deficits could occur in children exposed at levels too low to affect ChE activity. In December, 2014, EPA released a revised risk assessment for chlorpyrifos. The updated protocol continued the use of 10% RBC-ChE inhibition as the endpoint for cholinergic effects, but also considered neurodevelopmental outcomes measured in the prospective cohort study at Columbia Center for Children's Environmental Health (CCCEH) (Rauh, 2006). Fetal exposures in the CCCEH cohort were assumed to have resulted primarily from maternal contact with indoor crack and crevice treatment residue until that use for chlorpyrifos was phased out in December, 2001. For the 2014 risk assessment, EPA decided that a time-weighted average exposure via cord blood provided the most meaningful estimate of fetal exposure, and used physiologically based pharmacokinetic/ pharmacodynamic (PBPK/PD) modeling to estimate the time-weighted

average cord blood concentrations associated with adverse effects in the CCCEH study. In its November, 2016 human health risk assessment, EPA approved the use of an LOAEL for neurodevelopmental effects to act as the PoD for its chlorpyrifos RfD, as it had been satisfactorily shown that those effects emerge at levels of exposure too low to affect ChE activity. Because PBPK/PD models account for differing responses between humans and animals, the inter-species UF was reduced to 1. An SF of 10 was retained to adjust for the use of an LOAEL rather than an NOAEL as the PoD, and the intra-species UF of 10 was held at 10 to accommodate the increased sensitivity of children and pregnant mothers. The announcement of a new RfD awaits a final determination of fetal time-weighted average exposure, once dietary pesticide sources have been incorporated in the models (US EPA, 2016a).

## 5. Conclusions

Between February 21, 1986 and June 8, 2000, EPA's reference dose determinations for chlorpyrifos took their points of departure from the DOWCO 179 safety evaluation study's observations of plasma ChE inhibition. The original reported NOEL of 30µg/kg/day for repeated dietary exposure, when reduced by the intra-species UF of 10, established the first chronic dietary RfD at 3µg/kg/day. The chronic dietary RfD was reduced by an added factor of 10 to 0.3µg/kg/day following passage of FQPA in 1996. The new acute dietary RfD of 1µg/kg/day took its PoD from the DOWCO 179 study's finding of no significant effect on plasma ChE after 1 and 3 days of treatment at 100µg/kg/day.

Though the DOWCO 179 safety evaluation stands out as one of a very small number of controlled human studies to measure the toxicological effects of known exposures to chlorpyrifos, aspects of its realized design have detracted from its utility. Notably, treatment periods for the different dose groups did not all start and end on the same days; nor did the treatment periods last for equal lengths of time. The lack of synchrony and differences in duration between treatment periods limited the range of analyses that could be run using the tools available for analyzing longitudinal data in the early 1970's. We have reexamined the data using both the original method and a computational technique that accommodates the study's unbalanced design. Our findings, using both the original method of repeated measures ANOVA and LMM regression, disagree with the original determination of no response at the 30µg/kg/day level.

In our initial attempt to replicate the assessment by repeated measures ANOVA, we discovered that the original analyst had omitted one of two days' baseline data when testing subjects treated at 30μg/kg/day, yet had retained both baseline days when assessing the other two dose groups. Omission of those baseline measurements dramatically changes the results and interpretation of the ANOVA. Without them, there is no apparent treatment effect at the 30μg/kg/day dose level. In contrast, repeating the analysis with all baseline data gives a markedly different result from that originally reported: plasma-ChE inhibition for 30μg/kg/day treatment that is significant at p=0.012.

Our second analysis used LMM regression to assess dose-related trends in the treatment data. This method tolerates unbalanced data, and is therefore able to make use of measurements that could not be included in the original analysis by repeated measures ANOVA. We have considered two treatment models: one allows the dose-specific treatment effect estimates (slopes) to take on any value, and the second estimates a single treatment effect coefficient that scales linearly to all doses. The former (unconstrained) model estimates the 30μg/kg/day treatment effect as significantly different from zero at a rate of -0.070μmoles acetate/min/ml/day (95% CI -0.100, -0.040), and the 14μg/kg/day treatment effect as nearly significant at -0.029μmoles acetate/min/ml/day (95% CI -0.059, 0.002). The second (linearly constrained) model estimates significant non-zero treatment effects at all dose levels with dose-specific rates that are consistent with the unconstrained model estimates. Both models show a clear increase in treatment effect with increasing dose, and the similarity in slope estimates between the models suggests that the

assumption of an approximately linear relationship between dose and rate of change in plasma ChE activity is a reasonable assumption for these dose levels.

Our recovery phase analysis considered only the 30µg/kg/day and 100µg/kg/day dose groups, as each of the other groups had only two days of recovery data to support a slope estimate. Using LMM regression with linear splines, we found a surprising relationship between dose level and rate of recovery from plasma ChE inhibition. Subjects whose plasma ChE activity was more severely reduced by treatment at 100µg/kg/day recovered at a faster estimated rate than those whose activity was less inhibited by treatment at 30µg/kg/day. Slope estimates were 0.99µmoles acetate/min/ml/day (95% CI 0.008, 0.190) and 0.046µmoles acetate/min/ml/day (95% CI 0.007, 0.084), respectively. While a cautious interpretation is warranted because they were not statistically different from each other, the differing mean rates of recovery could reflect more aggressive recruitment of compensatory mechanisms in subjects whose ChE activity has been critically challenged.

In June of 2000, EPA revised its RfDs for chlorpyrifos. The updated values took their PoD from animal study data, and the determination considered RBC-ChE inhibition and developmental neurotoxicity along with plasma ChE inhibition as outcomes of interest. The revisions reduced the previous RfD by a factor 10 to protect infants and children from risk of adverse neurodevelopmental effects.

At the time, because EPA had imposed a temporary ban on the use of data from industry-sponsored pesticide studies on humans, the Coulston study was inadmissible as a

basis for the new RfDs. Between February, 1986 and June, 2000, however, correct interpretation of the DOWCO 179 data would have reduced the chlorpyrifos RfDs that were in effect. At a minimum, correct identification of the treatment effect at 30μg/kg/day, even using the original method, would have reduced the 1986 chronic dietary RfD from 3μg/kg/day to 1.4μg/kg/day, and the 1998 chronic dietary RfD from 0.3μg/kg/day to 0.14μg/kg/day.

While the estimated treatment effect at 14μg/kg/day does not quite show a significant difference from zero slope in our unconstrained model, the estimate provided by the linearly constrained model is unambiguously less than zero. That is, according to the linearly constrained model, the DOWCO 179 study did not identify an NOEL for plasma cholinesterase inhibition. When a study of otherwise high quality cannot establish an NOEL, EPA may use the LOEL as a point of departure and reduce it by an uncertainty factor up to 10 to compensate for the substitution (US EPA, 2015). Alternatively, at its discretion, it may refuse to set a tolerance. Either action would have reduced the tolerance for chlorpyrifos well below the RfD in effect at the time.

Given that credible epidemiologic evidence from the 1990's and early 2000's has associated adverse neurodevelopmental outcomes with prenatal exposures that were typical for the period, an earlier reduction in RfD and related exposure mitigation could have potentially reduced the incidence of adverse health effects in children of that era (Whyatt, 2004; Rauh, 2006; Rauh, 2011). It is regrettable that a hidden omission of valid

data from the analysis of this study's results may have adversely impacted public health for more than a decade.

**References**


Casarett, L.J., Klaassen, C.D., & Watkins, J.B. (2010). *Casarett & Doull's essentials of toxicology* (2nd ed.). New York, N.Y: McGraw-Hill Education LLC.

Christensen, K., Harper, B.; Luukinen, B., Buhl, K., Stone, D. (2009). *Chlorpyrifos Technical Fact Sheet*. National Pesticide Information Center, Oregon State University Extension Services. Retrieved August 7, 2017 from http://npic.orst.edu/factsheets/archive/chlorptech.html.

Coulston, F., Golberg, L., and Griffin, T. (1972). Safety evaluation of Dowco 179 in human volunteers. Unpublished report, the Institute of Experimental Pathology and Toxicology, Albany Medical College, Albany, NY.

Dow (The Dow Chemical Company). (2017). Dow statement on Associated Press story regarding chlorpyrifos. Retrieved November 13, 2017 from https://www.dow.com/en-us/news/press-releases/Dow%20ap%20Chlorpyrifos

EWG (Environmental Working Group). (1998). The English Patients: Human Experiments and Pesticide Policy. Retrieved January 30, 2018 from https://www.ewg.org/research/english-patients#.WpRb6MaZORs

Fitzmaurice, G., Laird, N.M., & Ware, J.H. (2004). *Applied longitudinal analysis* (Wiley series in probability and statistics). Hoboken, N.J.: Wiley-Interscience.

Griffin, P., Mason, H., Heywood, K., & Cocker, J. (1999). Oral and dermal absorption of chlorpyrifos: a human volunteer study. *Occupational and environmental medicine*, *56*(1), 10-13.

Kisicki, J. C., Seip, C. W., & Combs, M. L. (1999). A rising dose toxicology study to determine the No-Observable-Effect-Levels (NOEL) for erythrocyte acetylcholinesterase (AChE) inhibition and cholinergic signs and symptoms of chlorpyrifos at three dose levels. Unpublished report, MDC Harris Laboratory, Lincoln, NE.

Laird, N. M., & Ware, J. H. (1982). Random-effects models for longitudinal data. *Biometrics*, 963-974.

Nabb, D. P., & Whitfield, F. (1967). Determination of cholinesterase by an automated pH stat method. *Archives of Environmental Health: An International Journal*, *15*(2), 147-154.

Nolan, R. J., Rick, D. L., Freshour, N. L., & Saunders, J. H. (1984). Chlorpyrifos: pharmacokinetics in human volunteers. *Toxicology and applied pharmacology*, *73*(1), 8-15.

NRDC/PANNA (Natural Resources Defense Council & Pesticide Action Network of North America). (2007). Petition to revoke all tolerances and cancel all registrations for the pesticide chlorpyrifos. Retrieved January 30, 2018 from https://www.nrdc.org/sites/default/files/hea_10072201a.pdf

NRDC (Natural Resources Defense Council). (2017). Objections of the states of New York, Washington, California, Massachusetts, Maine, Maryland, and Vermont to EPA's March 29, 2017 order denying petition to revoke tolerances for chlorpyrifos and leaving tolerances in effect. Retrieved March 14, 2018 from https://www.nrdc.org/sites/default/files/seven-states-call-for-epa-to-ban-chlorpyrifos.pdf

Oak Ridge (Oak Ridge National Laboratory). (2009). Data Evaluation Record - Chlorpyrifos - Nonguideline - Study Type: Repeat Dose Oral Toxicity - Human. Toxicology and Hazard Assessment Group, Environmental Sciences Division. Oak Ridge, TN. May 26.

Oleskey, C., Fleischman, A., Goldman, L., Hirschhorn, K., Landrigan, P. J., Lappé, M., et al. (2004). Pesticide testing in humans: ethics and public policy. *Environmental Health Perspectives*, *112*(8), 914.

Park, C. (1972) Analysis of Plasma and Red Blood Cell - Cholinesterase Levels in Humans Ingesting DOWCO 179. Unpublished report, The Dow Chemical Company, Mathematical Applications Division, Computation Research Laboratory, Midland, MI.

Rauh, V. A., Garfinkel, R., Perera, F. P., Andrews, H. F., Hoepner, L., Barr, D. B., et al. (2006). Impact of prenatal chlorpyrifos exposure on neurodevelopment in the first 3 years of life among inner-city children. *Pediatrics*, *118*(6), e1845-e1859.

Rauh, V., Arunajadai, S., Horton, M., Perera, F., Hoepner, L., Barr, D.B., et al. (2011). Seven-year neurodevelopmental scores and prenatal exposure to chlorpyrifos, a common agricultural pesticide. *Environmental health perspectives*, *119*(8), 1196.

Resnik, D. B., & Portier, C. (2005). Pesticide testing on human subjects: weighing benefits and risks. *Environmental Health Perspectives*, *113*(7), 813.

Stefanidou, M., Athanaselis, S., & Spiliopoulou, H. (2009). Butyrylcholinesterase: biomarker for exposure to organophosphorus insecticides. *Internal medicine journal*, 39(1), 57-60.

Tang, J., Cao, Y., Rose, R. L., Brimfield, A. A., Dai, D., Goldstein, J. A., et al. (2001). Metabolism of chlorpyrifos by human cytochrome P450 isoforms and human, mouse, and rat liver microsomes. *Drug metabolism and disposition*, *29*(9), 1201-1204.

US EPA (United States Environmental Protection Agency). (1998a). Chlorpyrifos - FQPA Requirement - Report of the Hazard Identification Assessment Review Committee. Office of Pesticide Programs. Washington, DC, February 2.

US EPA (United States Environmental Protection Agency). (1998b). Chlorpyrifos - Re-evaluation - Report of the Hazard Identification Assessment Review Committee. Office of Pesticide Programs. Washington, DC, December 7.

US EPA (United States Environmental Protection Agency). (2002). Determination of the Appropriate FQPA Safety Factor(s) in Tolerance Assessment. Office of Pesticide Programs. Washington, DC, February 28.

US EPA (United States Environmental Protection Agency). (2006). Reregistration Eligibility Decision for Chlorpyrifos. Office of Pesticide Programs. Washington, DC, July 31.

US EPA (United States Environmental Protection Agency). (2008). FIFRA Scientific Advisory Committee Minutes September 16-18, 2008 - Scientific Issues Regarding the Agency's Evaluation of the Toxicity Profile for Chlorpyrifos. Office of Pesticide Programs. Washington, DC, December 17.

US EPA (United States Environmental Protection Agency). (2009). EPA Human Studies Review Board Meeting Report. Washington, DC, June 24-25.

US EPA (United States Environmental Protection Agency). (2016a). Chlorpyrifos: Revised Human Health Risk Assessment for Registration Review. Office of Chemical Safety and Pollution Prevention. Washington, DC, November 3.

US EPA (United States Environmental Protection Agency). (2016b). About pesticide tolerances. Retrieved November 30, 2017 from https://www.epa.gov/pesticide-tolerances/about-pesticide-tolerances

US EPA (United States Environmental Protection Agency). (2017). Chlorpyrifos; Order Denying PANNA and NRDC's Petition to Revoke Tolerances. Retrieved November 30, 2017 from https://www.epa.gov/sites/production/files/2017-03/documents/chlorpyrifos3b_order_denying_panna_and_nrdc27s_petitition_to_revoke_tolerances.pdf

Whyatt, R. M., Rauh, V., Barr, D. B., Camann, D. E., Andrews, H. F., Garfinkel, R., et al. (2004). Prenatal insecticide exposures and birth weight and length among an urban minority cohort. *Environmental health perspectives*, *112*(10), 1125.

## Appendices

### Appendix A: original data set with handwritten notes (Park, 1972)

Annotations circled in green were used to define treatment periods for our analyses.

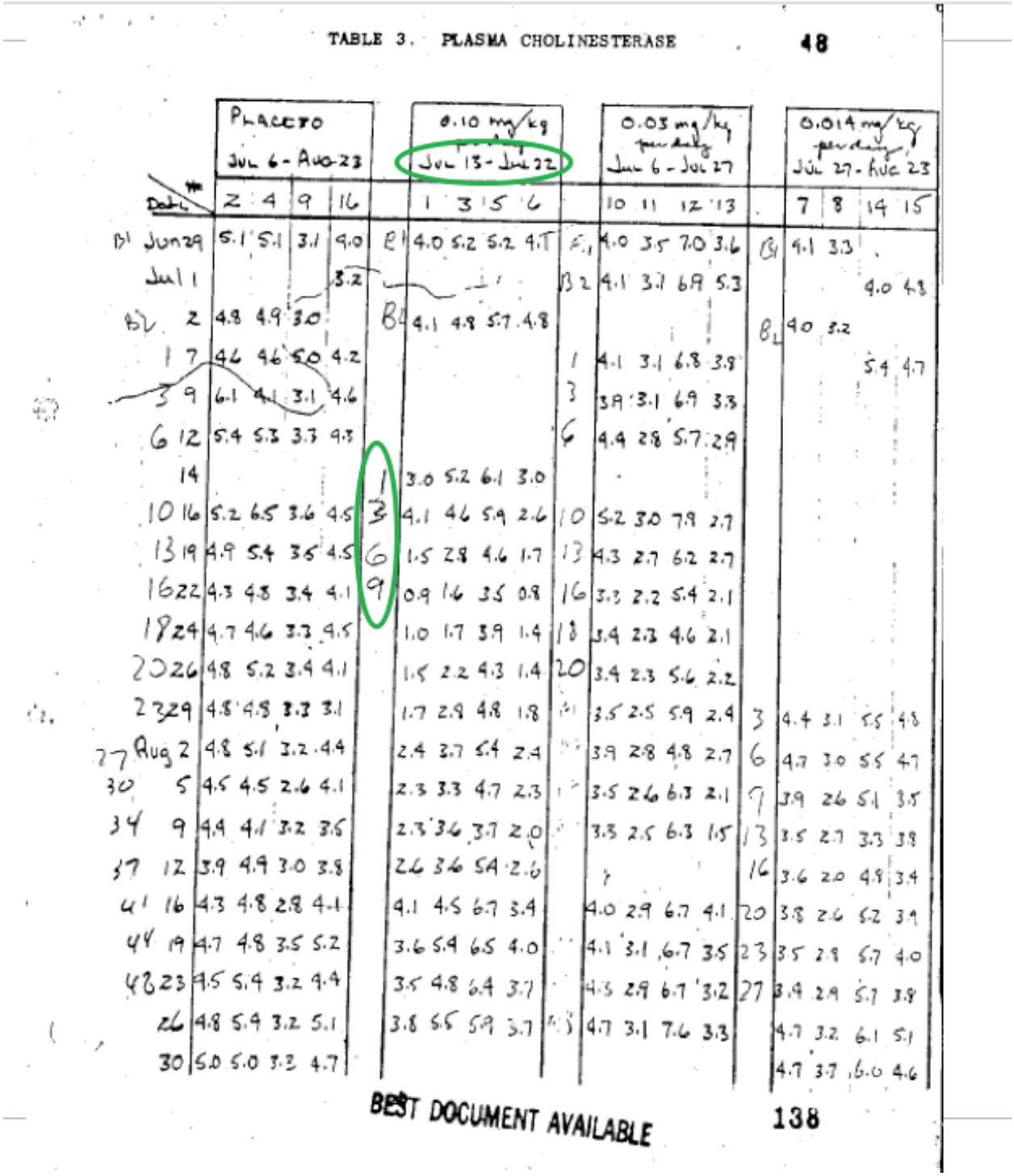


TABLE 3. PLASMA CHOLINESTERASE 48

| Date | Placebo Jul 6 - Aug 23 | | | | 0.10 mg/kg Jul 13 - Jul 22 | | | | 0.03 mg/kg per day Jun 6 - Jul 27 | | | | 0.014 mg/kg per day Jul 27 - Aug 23 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 4 | 9 | 16 | 1 | 3 | 5 | 6 | 10 | 11 | 12 | 13 | 7 | 8 | 14 | 15 |
| Jun 29 | 5.1 | 5.1 | 3.1 | 4.0 | 4.0 | 5.2 | 5.2 | 4.7 | 4.0 | 3.5 | 7.0 | 3.6 | 4.1 | 3.3 | | |
| Jul 1 | | | | 3.2 | | | | | 4.1 | 3.1 | 6.9 | 5.3 | | | 4.0 | 4.8 |
| 2 | 4.8 | 4.9 | 3.0 | | 4.1 | 4.8 | 5.7 | 4.8 | | | | | 4.0 | 3.2 | | |
| 7 | 4.6 | 4.6 | 5.0 | 4.2 | | | | | 4.1 | 3.1 | 6.8 | 3.8 | | | 5.4 | 4.7 |
| 9 | 6.1 | 4.1 | 3.1 | 4.6 | | | | | 3.9 | 3.1 | 6.9 | 3.3 | | | | |
| 12 | 5.4 | 5.3 | 3.7 | 4.3 | | | | | 4.4 | 2.8 | 5.7 | 2.9 | | | | |
| 14 | | | | | 3.0 | 5.2 | 6.1 | 3.0 | | | | | | | | |
| 16 | 5.2 | 6.5 | 3.6 | 4.5 | 4.1 | 4.6 | 5.9 | 2.6 | 5.2 | 3.0 | 7.9 | 2.7 | | | | |
| 19 | 4.9 | 5.4 | 3.5 | 4.5 | 1.5 | 2.8 | 4.6 | 1.7 | 4.3 | 2.7 | 6.2 | 2.7 | | | | |
| 22 | 4.3 | 4.8 | 3.4 | 4.1 | 0.9 | 1.6 | 3.5 | 0.8 | 3.3 | 2.2 | 5.4 | 2.1 | | | | |
| 24 | 4.7 | 4.6 | 3.3 | 4.5 | 1.0 | 1.7 | 3.9 | 1.4 | 3.4 | 2.3 | 4.6 | 2.1 | | | | |
| 26 | 4.8 | 5.2 | 3.4 | 4.1 | 1.5 | 2.2 | 4.3 | 1.4 | 3.4 | 2.3 | 5.6 | 2.2 | | | | |
| 29 | 4.8 | 4.8 | 3.3 | 3.1 | 1.7 | 2.9 | 4.8 | 1.8 | 3.5 | 2.5 | 5.9 | 2.4 | 4.4 | 3.1 | 5.5 | 4.8 |
| Aug 2 | 4.8 | 5.1 | 3.2 | 4.4 | 2.4 | 3.7 | 5.4 | 2.4 | 3.9 | 2.8 | 4.8 | 2.7 | 4.7 | 3.0 | 5.5 | 4.7 |
| 5 | 4.5 | 4.5 | 2.6 | 4.1 | 2.3 | 3.3 | 4.7 | 2.3 | 3.5 | 2.6 | 6.3 | 2.1 | 3.9 | 2.6 | 5.1 | 3.5 |
| 9 | 4.4 | 4.1 | 3.2 | 3.5 | 2.3 | 3.6 | 3.7 | 2.0 | 3.3 | 2.5 | 6.3 | 1.5 | 3.5 | 2.7 | 3.3 | 3.8 |
| 12 | 3.9 | 4.9 | 3.0 | 3.8 | 2.6 | 3.6 | 5.4 | 2.6 | | | | | 3.6 | 2.0 | 4.9 | 3.4 |
| 16 | 4.3 | 4.8 | 2.8 | 4.1 | 4.1 | 4.5 | 6.7 | 3.4 | 4.0 | 2.9 | 6.7 | 4.1 | 3.8 | 2.6 | 5.2 | 3.9 |
| 19 | 4.7 | 4.8 | 3.5 | 5.2 | 3.6 | 5.4 | 6.5 | 4.0 | 4.1 | 3.1 | 6.7 | 3.5 | 3.5 | 2.9 | 5.7 | 4.0 |
| 23 | 4.5 | 5.4 | 3.2 | 4.4 | 3.5 | 4.8 | 6.4 | 3.7 | 4.3 | 2.9 | 6.7 | 3.2 | 3.9 | 2.9 | 5.7 | 3.8 |
| 26 | 4.8 | 5.4 | 3.2 | 5.1 | 3.8 | 6.5 | 5.9 | 3.7 | 4.7 | 3.1 | 7.6 | 3.3 | 4.7 | 3.2 | 6.1 | 5.1 |
| 30 | 5.0 | 5.0 | 3.2 | 4.7 | | | | | | | | | 4.7 | 3.7 | 6.0 | 4.6 |



138

**Appendix B: model output for treatment analysis by repeated measures ANOVA**

**Original ANOVA results reported by Park:**

**a) for 14µg/kg/day dose group**

PLASMA

ANALYSIS OF VARIANCE FOR 0.014 mg/kg/day

| Source | SS | d.f. | M.S. | F |
|---|---|---|---|---|
| Between Subjects | 42.396 | 7 | | |
| - dose level | 0.435 | 1 | 0.435 | |
| - error | 41.961 | 6 | 6.994 | |
| Within Subjects | 17.718 | 72 | | |
| - days | 6.140 | 9 | 0.682 | |
| - days x dose | 1.726 | 9 | 0.192 | 1.05 |
| - error | 9.853 | 54 | 0.182 | |
| TOTAL | 60.114 | 79 | | |

F for homogeneity of variances = 1.45

**b) for 30µg/kg/day dose group**

PLASMA

ANALYSIS OF VARIANCE FOR 0.03 mg/kg/day

| Source | SS | d.f. | M.S. | F |
|---|---|---|---|---|
| Between Subjects | 89.360 | 7 | | |
| - dose levels | 4.909 | 1 | 4.909 | |
| - error | 84.451 | 6 | 14.075 | |
| Within Subjects | 24.560 | 64 | | |
| - days | 9.535 | 8 | 1.192 | |
| - days x dose | 3.311 | 8 | 0.414 | 1.70 |
| - error | 11.714 | 48 | 0.244 | |
| TOTAL | 113.920 | 71 | | |

F for homogeneity of variances = 1.59

**c) for 100µg/kg/day dose group**

PLASMA

ANALYSIS OF VARIANCE FOR CONTROL VS. 0.10 mg/kg/day

| Source | Mean Square | d.f. | Sums of Squares | F |
|---|---|---|---|---|
| Between Subjects | 33.036 | 7 | | |
| - dose level | 5.929 | 1 | 5.929 | |
| - error | 27.107 | 6 | 4.518 | |
| Within Subjects | 37.959 | 32 | | |
| - days | 16.733 | 4 | 4.183 | |
| - days x dose | 16.043 | 4 | 4.011 | 18.57* |
| - error | 5.183 | 24 | 0.216 | |
| TOTAL | 70.996 | 41 | | |

F for homogeneity of variances:

$F_{2,22} = 2.55$

* Significant at $\alpha = 0.05$

**Updated results:**

F values and probabilities highlighted in yellow correspond to those reported in Park's original results summary above. Those highlighted in green differ from the original findings due to the inclusion of baseline measurements that were omitted from the original analysis.

**a) for 14µg/kg/day dose group**

```
Error: fsubj
          Df Sum Sq Mean Sq F value Pr(>F)
fdose      1   0.44   0.435   0.062  0.811
Residuals  6  41.96   6.993

Error: Within
            Df Sum Sq Mean Sq F value  Pr(>F)
fdate        9  6.140  0.6822   3.739 0.00105 **
fdose:fdate  9  1.726  0.1918   1.051 0.41312
Residuals   54  9.853  0.1825
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

**b) for 30μg/kg/day dose group**

**i) using original baseline data (i.e., omitting second day)**

```
Error: fsubj
          Df Sum Sq Mean Sq F value Pr(>F)
fdose      1   4.91   4.909   0.349  0.576
Residuals  6  84.45  14.075

Error: fsubj:fdate
            Df Sum Sq Mean Sq F value   Pr(>F)
fdate        8  9.535  1.1919   4.884 0.000191 ***
fdose:fdate  8  3.311  0.4139   1.696 0.123785
Residuals   48 11.714  0.2440
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

**ii) using all baseline data (i.e., including second day)**

```
Error: fsubj
          Df Sum Sq Mean Sq F value Pr(>F)
fdose      1   2.93   2.926   0.191  0.677
Residuals  6  91.90  15.316

Error: Within
            Df Sum Sq Mean Sq F value   Pr(>F)
fdate        9  9.811  1.0901   3.826 0.000862 ***
fdose:fdate  9  6.825  0.7583   2.661 0.012432 *
Residuals   54 15.387  0.2849
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

**c) for 100µg/kg/day dose group**

```
Error: fsubj
          Df Sum Sq Mean Sq F value Pr(>F)
fdose      1  5.929   5.929   1.312  0.296
Residuals  6 27.107   4.518

Error: Within
            Df Sum Sq Mean Sq F value   Pr(>F)
fdate        4 16.734   4.183   19.37 3.11e-07 ***
fdose:fdate  4 16.044   4.011   18.57 4.52e-07 ***
Residuals   24  5.183   0.216
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

## Appendix C: model output for treatment analysis by linear mixed effects regression

### 1. Treatment model with categorical dose predictor

```
Linear mixed model fit by REML t-tests use Satterthwaite approximations
  to degrees of freedom [lmerMod]
Formula: plasma ~ fdose * dft + fnewd + (1 | fsubj)
   Data: d

REML criterion at convergence: 332.6

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.5170 -0.5582  0.0175  0.4207  3.6756

Random effects:
 Groups   Name        Variance Std.Dev.
 fsubj    (Intercept) 1.2368   1.1121
 Residual             0.2344   0.4842
Number of obs: 176, groups:  fsubj, 16

Fixed effects:
              Estimate Std. Error        df t value Pr(>|t|)
(Intercept)    4.32261    0.58124  13.74000   7.437 3.55e-06 ***
fdose14       -0.14122    0.81631  13.36000  -0.173 0.865244
fdose30        0.24249    0.80569  12.68000   0.301 0.768313
fdose100       0.32401    0.81520  13.29000   0.397 0.697341
dft           -0.01721    0.01630 136.98000  -1.056 0.292840
fnewd2        -0.08792    0.23949 137.16000  -0.367 0.714104
fnewd3         0.05029    0.21401 137.00000   0.235 0.814562
fnewd8         0.16044    0.21064 137.11000   0.762 0.447550
fnewd10        0.10040    0.22550 137.01000   0.445 0.656843
fnewd13        0.13215    0.23454 137.00000   0.563 0.574049
fnewd15        0.04429    0.30093 136.98000   0.147 0.883197
fnewd17        0.84448    0.22981 136.99000   3.675 0.000341 ***
fnewd20        0.40655    0.27343 136.99000   1.487 0.139355
fnewd23        0.17696    0.32657 136.99000   0.542 0.588773
fnewd25        0.18417    0.34793 136.98000   0.529 0.597440
fnewd27        0.47617    0.37382 136.98000   1.274 0.204901
fnewd30        0.23959    0.33860 136.99000   0.708 0.480394
fnewd34        0.54268    0.39515 136.99000   1.373 0.171879
fnewd37        0.06217    0.44992 136.98000   0.138 0.890303
fnewd41       -0.09935    0.52727 136.98000  -0.188 0.850822
fnewd44        0.04514    0.58744 136.98000   0.077 0.938869
fnewd48        0.49612    0.66962 136.98000   0.741 0.460031
fnewd51        0.92810    0.73231 136.98000   1.267 0.207175
fnewd55        0.94158    0.81689 136.98000   1.153 0.251063
fdose14:dft   -0.02857    0.01542 136.98000  -1.853 0.065997 .
fdose30:dft   -0.07008    0.01514 136.99000  -4.628 8.46e-06 ***
fdose100:dft  -0.34870    0.04084 137.00000  -8.538 2.31e-14 **
```

## 2. Treatment model with continuous dose predictor

```
Linear mixed model fit by REML t-tests use Satterthwaite approximations
  to degrees of freedom [lmerMod]
Formula: plasma ~ dose * dft + fnewd + (1 | fsubj)
   Data: d

REML criterion at convergence: 346.4

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.5018 -0.5636 -0.0238  0.4660  3.5421

Random effects:
 Groups   Name        Variance Std.Dev.
 fsubj    (Intercept) 1.1027   1.0501
 Residual             0.2403   0.4902
Number of obs: 176, groups:  fsubj, 16

Fixed effects:
             Estimate Std. Error         df t value Pr(>|t|)
(Intercept)  4.401e+00  3.904e-01  1.843e+01  11.275 1.05e-09 ***
dose         1.162e-03  7.074e-03  1.533e+01   0.164   0.8717
dft         -1.940e-02  1.148e-02  1.516e+02  -1.690   0.0930 .
fnewd2      -2.166e-01  2.350e-01  1.395e+02  -0.922   0.3583
fnewd3       1.249e-01  2.133e-01  1.396e+02   0.586   0.5592
fnewd8       8.515e-02  2.105e-01  1.392e+02   0.405   0.6865
fnewd10      5.848e-02  2.260e-01  1.393e+02   0.259   0.7962
fnewd13      1.255e-01  2.324e-01  1.392e+02   0.540   0.5900
fnewd15      1.950e-01  2.968e-01  1.393e+02   0.657   0.5122
fnewd17      8.890e-01  2.205e-01  1.405e+02   4.031 9.07e-05 ***
fnewd20      4.256e-01  2.556e-01  1.417e+02   1.665   0.0982 .
fnewd23      1.705e-01  2.995e-01  1.427e+02   0.569   0.5700
fnewd25      3.188e-01  3.137e-01  1.432e+02   1.016   0.3113
fnewd27      6.343e-01  3.330e-01  1.438e+02   1.905   0.0588 .
fnewd30      2.439e-01  3.000e-01  1.469e+02   0.813   0.4177
fnewd34      5.742e-01  3.291e-01  1.480e+02   1.745   0.0831 .
fnewd37      1.199e-01  3.565e-01  1.487e+02   0.336   0.7371
fnewd41     -6.728e-03  3.952e-01  1.493e+02  -0.017   0.9864
fnewd44      1.639e-01  4.257e-01  1.497e+02   0.385   0.7007
fnewd48      6.498e-01  4.676e-01  1.500e+02   1.390   0.1666
fnewd51      1.108e+00  4.998e-01  1.502e+02   2.217   0.0281 *
fnewd55      1.156e+00  5.436e-01  1.504e+02   2.127   0.0350 *
dose:dft    -2.975e-03  3.473e-04  1.408e+02  -8.565 1.71e-14 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

**Appendix D: model output for treatment and recovery analysis by linear mixed effects regression with splines**

1. Raw R output

```
Linear mixed model fit by REML t-tests use Satterthwaite approximations to
degrees of
  freedom [lmerMod]
Formula: plasma ~ 1 + fdose * dft + recov30 + recov100 + fnewd + (1 |
fsubj)
   Data: d

REML criterion at convergence: 423.3

Scaled residuals:
    Min      1Q  Median      3Q     Max
-2.6517 -0.4834  0.0274  0.4563  3.9947

Random effects:
 Groups   Name        Variance Std.Dev.
 fsubj    (Intercept) 1.3100   1.1446
 Residual             0.2128   0.4613
Number of obs: 236, groups:  fsubj, 16

Fixed effects:
              Estimate Std. Error         df t value Pr(>|t|)
(Intercept)    4.337400   0.594449  13.660000   7.297 4.54e-06 ***
fdose14       -0.148180   0.835639  13.330000  -0.177 0.861918
fdose30        0.223220   0.825723  12.710000   0.270 0.791243
fdose100       0.296201   0.834410  13.260000   0.355 0.728187
dft           -0.009152   0.014671 195.010000  -0.624 0.533492
recov30        0.116272   0.021563 195.020000   5.392 2.00e-07 ***
recov100       0.421190   0.035350 195.030000  11.915  < 2e-16 ***
fnewd2        -0.090397   0.227584 195.230000  -0.397 0.691653
fnewd3         0.049934   0.203641 195.040000   0.245 0.806554
fnewd8         0.145275   0.200382 195.160000   0.725 0.469326
fnewd10        0.061565   0.213582 195.050000   0.288 0.773460
fnewd13        0.059641   0.218789 195.040000   0.273 0.785452
fnewd15        0.012000   0.285820 195.010000   0.042 0.966553
fnewd17        0.725229   0.206176 195.030000   3.518 0.000542 ***
fnewd20        0.219548   0.234231 195.030000   0.937 0.349757
fnewd23       -0.077800   0.269786 195.030000  -0.288 0.773365
fnewd25        0.001508   0.287787 195.030000   0.005 0.995823
fnewd27        0.230817   0.307346 195.020000   0.751 0.453559
fnewd30        0.220765   0.307594 195.020000   0.718 0.473793
fnewd34        0.375769   0.358285 195.020000   1.049 0.295568
fnewd37       -0.054274   0.403486 195.020000  -0.135 0.893136
fnewd41       -0.423498   0.467026 195.020000  -0.907 0.365634
fnewd44       -0.274721   0.519556 195.020000  -0.529 0.597573
fnewd48        0.277236   0.583751 195.010000   0.475 0.635375
fnewd51        0.397193   0.635141 195.010000   0.625 0.532465
fnewd55        0.497995   0.726708 195.010000   0.685 0.493984
fdose14:dft   -0.018934   0.013349 195.010000  -1.418 0.157652
fdose30:dft   -0.063747   0.012513 195.020000  -5.095 8.22e-07 ***
fdose100:dft  -0.311450   0.028782 195.040000 -10.821  < 2e-16 ***
```

**Appendix E: R scripts used in the analyses**

All require access to the data file 'toxdata.csv'.

Variables in 'toxdata.csv':

plasma = plasma cholinesterase activity measurements in µmoles acetate/min/ml
rbc = red blood cell cholinesterase activity measurements in µmoles acetate/min/ml
dose = dose level in µg/kg/day (levels: 0, 14, 30, 100)
subject = subject ID number (integer, 1 to 16)
newdate = count of days since first baseline measurement (integer, 0 to 62)
daysfromtrt = count of days since first treatment (integer, 0 to 55)
trt = binary indicator of treatment period measurement
recover = binary indicator of recovery period measurement
date = calendar date
visit = alphanumeric code indicating category of measurement and index of day (B = baseline, P = placebo, D = non-zero dose goup, R = recovery)

**1a. Replication of original ANOVA: 14µg/kg/day vs. control**

```
### Replication of original ANOVA: 14µg/kg/day vs. control
### Degrees of freedom match original
### Mean squares match original

rm(list=ls())
setwd("Users/seth/Desktop/Thesis/Analysis/ANOVA")

d0 <- read.csv("/Users/seth/Desktop/Thesis/Analysis/DataAndScripts/toxdata.csv",header=T)

## simplify table by removing rbc and recover columns
d1 <- d0[,!(names(d0) %in% c("rbc","recover"))]

## remove recovery phase rows
d <- d1[!grepl("R",d1$visit),]

d14 <- d[(d$dose==0 | d$dose==14),]

d14$dft <- d14$daysfromtrt
d14$dft <- ifelse(d14$dft==10,9,d14$dft)

### Recode newdate to treat all baseline data as balanced and taken
### on exactly 2 occasions (newdates) per subject

d14$newdate <- ifelse(d14$newdate==0 & d14$subj==16,-1,d14$newdate)
```

```
d14$newdate <- ifelse(d14$newdate==2 & d14$subj==16,0,d14$newdate)
d14$newdate <- ifelse(d14$newdate==8 & d14$subj==16,3,d14$newdate)
d14$newdate <- ifelse(d14$newdate==2 & d14$subj %in% c(14,15),0,d14$newdate)
d14$newdate <- ifelse(d14$newdate==8 & d14$subj %in% c(14,15),3,d14$newdate)

newdate14 <- c(0,3,30,34,37,41,44,48,51,55)
d14_bal <- d14[d14$newdate %in% newdate14,]

attach(d14_bal)

fdose <- as.factor(dose)
fsubj <- as.factor(subject)
fdate <- as.factor(newdate)
aov14 <- aov(plasma ~ fdose*fdate + Error(fsubj),data=d14_bal)

summary(aov14)
```

**1b. Replication of original ANOVA: 30µg/kg/day vs. control**

```
### Replication of original ANOVA: 30µg/kg/day vs. control
### Degrees of freedom match original
### Mean squares match original
### No significant difference between 30µg/kg/day and controls

rm(list=ls())
setwd("Users/seth/Desktop/Thesis/Analysis/ANOVA")

d0 <- read.csv("/Users/seth/Desktop/Thesis/Analysis/DataAndScripts/toxdata.csv",header=T)

## simplify table by removing rbc and recover columns
d1 <- d0[,!(names(d0) %in% c("rbc","recover"))]

## remove recovery phase rows
d <- d1[!grepl("R",d1$visit),]

d30 <- d[(d$dose==0 | d$dose==30),]

d30$newdate <- ifelse(d30$newdate==3,2,d30$newdate)
newdate30 <- c(0,2,8,10,13,17,20,23,25,27)

## remove second day of baseline data
d30_park <- d30[d30$newdate %in% newdate30[-2],]

attach(d30_park)
```

```
fdose <- as.factor(dose)
fsubj <- as.factor(subject)
fdate <- as.factor(newdate)
aov30_park <- aov(plasma ~ fdose*fdate + Error(fsubj/fdate), data = d30_park)

summary(aov30_park)
```

**1c. Updated ANOVA: 30µg/kg/day vs. control**

```
### Degrees of freedom do not match original
### Mean squares do not match original
### Significant difference between 30µg/kg/day and controls

rm(list=ls())
setwd("Users/seth/Desktop/Thesis/Analysis/ANOVA")

d0 <- read.csv("/Users/seth/Desktop/Thesis/Analysis/DataAndScripts/toxdata.csv",header=T)

## simplify table by removing rbc and recover columns
d1 <- d0[,!(names(d0) %in% c("rbc","recover"))]

## remove recovery phase rows
d <- d1[!grepl("R",d1$visit),]

d30 <- d[(d$dose==0 | d$dose==30),]

d30$newdate <- ifelse(d30$newdate==3,2,d30$newdate)
newdate30 <- c(0,2,8,10,13,17,20,23,25,27)

d30_full <- d30[d30$newdate %in% newdate30,]

attach(d30_full)

fdose <- as.factor(dose)
fsubj <- as.factor(subject)
fdate <- as.factor(newdate)
aov30_full <- aov(plasma ~ fdose*fdate + Error(fsubj), data = d30_full)

summary(aov30_full)
```

**1d. Replication of original ANOVA: 100µg/kg/day vs. control**

```
### Fully replicates original result with same number of degrees of freedom

rm(list=ls())
setwd("Users/seth/Desktop/Thesis/Analysis/ANOVA")

d0 <- read.csv("/Users/seth/Desktop/Thesis/Analysis/DataAndScripts/toxdata.csv",header=T)

## simplify table by removing rbc and recover columns
d1 <- d0[,!(names(d0) %in% c("rbc","recover"))]

## remove recovery phase rows
d <- d1[!grepl("R",d1$visit),]

d100 <- d[(d$dose==0 | d$dose==100),]

d100$dft <- d100$daysfromtrt

### enhance balance of data set by defining days 9 and 10
### as equivalent time points
d100$dft <- ifelse(d100$dft==10,9,d100$dft)

### enhance balance of data set by defining days 2 and 3
### as equivalent time points (only subject 16 had a baseline
### measurement recorded on newdate = 2)
d100$newdate <- ifelse(d100$newdate==2,3,d100$newdate)

dft100 <- c(0,3,6,9)
dft0 <- c(0,9,13,16)
d100_bal <- d100[(d100$dft %in% dft100 & d100$dose==100)|
                 (d100$dft %in% dft0 & d100$dose==0),]

attach(d100_bal)

fdose <- as.factor(dose)
fsubj <- as.factor(subject)
fdate <- as.factor(newdate)
aov100 <- aov(plasma ~ fdose*fdate + Error(fsubj),data=d100_bal)

summary(aov100)
```

**2. Treatment effect analysis by linear mixed effects regression**

```
################################################################################
####
#### THIS SCRIPT ENCODES TREATMENT MODELS 1 AND 2
#### MODEL 1: dose in dose*day is a factor variable
#### MODEL 2: dose in dose*day is continuous
####
################################################################################

## read in data from csv file
# 'toxdata.csv' is the complete set of ChE activity measurements for
# plasma and rbc ChE at all doses and for all treatment dates

setwd("~/Desktop/Thesis/Analysis/DataAndScripts")
d0 <- read.csv("toxdata.csv",header=T)

## remove recovery phase rows and rename working data file
d <- d0[!grepl("R",d0$visit),]

## option to write working data file to directory for proofreading

# write.csv(d,"d0minusrecovery.csv")

## load 'lme4' mixed effects model package
## load 'lmerTest' package to include significance test results in
## model summaries

library(lme4)
library(lmerTest)

## code dose, subject, and daysfromtrt (dft) as factors

d$fdose <- as.factor(d$dose)
d$fsubj <- as.factor(d$subject)
d$dft <- d$daysfromtrt
d$fdft <- as.factor(d$dft)
d$fnewd <- as.factor(d$newdate)

############################################################

### MODEL 1: TREATS SUBJECT AND STUDY DATE AS FACTORS.
### FOR BOTH INTERCEPT AND DOSE*DAY ESTIMATION, DOSE IS A FACTOR.
### DAY OF TREATMENT IS CONTINUOUS.

fitfull_fdose <- lmer(plasma ~ fdose*dft + fnewd + (1|fsubj), data=d)
```

```
summary(fitfull_fdose)


## If desired, run model without calendar date confounding variable,
## and compute likelihood ratio in ANOVA

# fitmin_fdose <- lmer(plasma ~ fdose*dft  + (1|fsubj), data=d)
# summary(fitmin_fdose)

# anova(fitfull_fdose,fitmin_fdose)

##########################################################

### MODEL 2: TREATS SUBJECT AND STUDY DATE AS FACTORS.
### FOR INTERCEPT AND DOSE*DAY
### ESTIMATION, DOSE IS CONTINUOUS.
### DAY OF TREATMENT IS CONTINUOUS.

fitfull_cdose <- lmer(plasma ~ dose*dft + fnewd + (1|fsubj), data=d)
summary(fitfull_cdose)


## If desired, run model without calendar date confounding, and compute
## likelihood ratio in ANOVA

# fitmin_cdose <- lmer(plasma ~ dose*dft + (1|fsubj), data=d)
# summary(fitmin_cdose)

# anova(fitfull_cdose,fitmin_cdose)

##########################################################
```

**3a. Treatment/recovery effect analysis by linear mixed affects regression with splines**

```
#########################################################################
#########################################################################
#### THIS IS A MODIFIABLE SCRIPT TO ALLOW CHANGES IN THE KNOT LOCATION
#### COMMENT LINES 99 TO 122 IN/OUT TO SELECT KNOT POSITION
####
####
#### THE ORIGINAL SCRIPT USES A SPLINED LINEAR MIXED EFFECTS MODEL
#### TO APPROXIMATE MEAN PChE LEVELS BY TREATMENT DAY AND
#### RECOVERY DAY, ADJUSTING FOR SUBJECT-SPECIFIC RANDOM
#### EFFECTS AND DATE-SPECIFIC TRENDS IN THE LABORATORY
#### ANALYSES. IT PLOTS THE ORIGINAL DATA ON FOUR
```

```
#### SEPARATE GRAPHS, ONE FOR EACH DOSE GROUP.
####
#########################################################################
#########################################################################


#############################################################
### PRELIMINARY STEPS 1 to 4 BELOW ONLY NEED TO RUN ONCE
#############################################################

## 1. SET WORKING DIRECTORY
setwd("~/Desktop/Thesis/Analysis/DataAndScripts")

## 2. LOAD REQUIRED MIXED EFFECTS REGRESSION PACKAGES
library(lme4)
library(lmerTest)

## 3. CLEAR ANY EXISTING VARIABLES FROM WORKING ENVIRONMENT
# rm(list=ls())

## 4. READ IN DATA
# 'toxdata.csv' is the complete set of ChE activity measurements for
# plasma and rbc ChE at all doses and for all study dates,
# including baseline, treatment, and recovery periods

d0 <- read.csv("toxdata.csv", header=T)


################ END PRELIMINARY STEPS #################



#########################################################
### **** BEGIN SECOND AND LATER EXECUTIONS HERE ****
#########################################################

## COPY 'MASTER' DATA FRAME d0 INTO A MODIFIABLE DATA FRAME d

d <- d0

## IF DESIRED, DROP LAST 2 RECOVERY DAYS, DAYS 55 AND 58, FROM
## 100µg GROUP DATE, PER RECOVERY SENSITIVITY ANALYSIS THAT FOUND
## RECOVERY SLOPE INCREASED TO A STABLE VALUE WHEN THOSE DATES
## WERE REMOVED; TO USE FULL DATSET, COMMENT OUT THE NEXT LINE
d <- subset(d,!(dose==100 & newdate > 51))
```

```
## IF DESIRED, DROP LAST 2 RECOVERY DAYS, DAYS 55 AND 58, FROM
## 30µg GROUP DATE, PER RECOVERY SENSITIVITY ANALYSIS. THOUGH
## RECOVERY SLOPE DOES NOT INCREASE TO A STABLE VALUE WHEN THOSE
DATES
## ARE REMOVED, MODEL FIT TO MEAN LINE IS IMPROVED (SEE RESIDUAL
PLOTS)
## TO USE FULL DATSET, COMMENT OUT THE NEXT LINE
d <- subset(d,!(dose==30 & newdate > 51))


## 112617 UPDATE DROPS THE LAST TWO DAYS FROM 14µg GROUP DATA
## AND CONTROL DATA
##
d <- subset(d,!(dose==14 & newdate > 55))
d <- subset(d,!(dose==0 & newdate > 55))

# DEFINE DOSE GROUP MEMBERSHIP BY SUBJECT ID
subs <- sort(unique(d$subject))
subs0 <- c(2,4,9,16)
subs14 <- c(7,8,14,15)
subs30 <- c(10,11,12,13)
subs100 <- c(1,3,5,6)




###=================================================================
=========
### I. THIS SECTION CREATES THE LINEAR SPLINES MODEL WITH A SINGLE KNOT
AT
### THE TRANSITION DAY BETWEEN TREATMENT AND RECOVERY PERIODS.
### SPECIFICALLY, THE KNOT IS THE TIME POINT LOCATED (APPROXIMATELY) 24
HOURS
### AFTER THE FINAL TREATMENT. IT IS BOTH THE LAST TREATMENT DAY AND
### RECOVERY DAY ZERO
###=================================================================
=========



## DEFINE END OF TREATMENT PERIOD FOR ALL DOSE GROUPS
## The 'lasttrt' variable stores the study calendar date for
## the last day of the each group's treatment period

doses <- c(0,14,30,100)
```

```
### FOLLOWING LINE SETS THE PREFERRED LOCATION FOR THE KNOT,
### ON THE LAST TREATMENT DAY
lasttrt <- c(55,27,20,9)

# FOLLOWING LINE SETS PENULTIMATE *TREATMENT* DAY OR
(EQUIVALENTLY)
# PREVIOUS DATA POINT AS KNOT POSITION FOR EACH GROUP
# NOTE: NUMBER OF DAYS ELAPSED SINCE LAST TREATMENT
# DIFFERS BETWEEN DOSE GROUPS lasttrt <- c(55,27,18,6)
# lasttrt <- c(55,27,18,6)

# FOLLOWING LINE SETS LAST TREATMENT MINUS 1 DAY AS KNOT POSITION
# lasttrt <- c(55,27,19,8)

# FOLLOWING LINE SETS LAST TREATMENT MINUS 2 DAYS AS KNOT POSITION
# lasttrt <- c(55,27,18,7)

# FOLLOWING LINE SETS LAST TREATMENT MINUS 3 DAYS AS KNOT POSITION
# lasttrt <- c(55,27,17,6)

# FOLLOWING LINE SETS LAST TREATMENT PLUS 1 DAY AS KNOT POSITION
# lasttrt <- c(55,27,21,10)

# FOLLOWING LINE SETS LAST TREATMENT PLUS 2 DAYS AS KNOT POSITION
# lasttrt <- c(55,27,22,11)


## Shorten predictor names, creating factor versions of dose, subject,
## and daysfromtrt (dft) variables.

d$dft[d$dose==0] <- d$daysfromtrt[d$dose==0]
d$dft[d$dose==14] <- d$daysfromtrt[d$dose==14]
d$dft[d$dose==30] <- d$daysfromtrt[d$dose==30]
d$dft[d$dose==100] <- d$daysfromtrt[d$dose==100]

d$fdose <- as.factor(d$dose)
d$fsubj <- as.factor(d$subject)
d$fdft <- as.factor(d$dft)
d$fnewd <- as.factor(d$newdate)


## OPTION: EXECUTE THE FOLLOWING LINE TO DROP RECOVERY DAYS FROM
14µg GROUP DATA
# d <- d[!(grepl("R",d0$visit)&d$dose==14),]
```

```
## OPTION: EXECUTE THE FOLLOWING LINE TO DROP RECOVERY DAYS FOR ALL
GROUPS
# d <- d[!(grepl("R",d0$visit)),]

## THE LINES BELOW ENCODE THE MODELS USED ON AUGUST 1ST --

d$recspline <- (d$dft - lasttrt[match(d$fdose, doses)]) * (d$dft > lasttrt[match(d$fdose, doses)])

d$recov14 <- ifelse(d$dose==14, 1, 0)*d$recspline
d$recov30 <- ifelse(d$dose==30, 1, 0)*d$recspline
d$recov100 <- ifelse(d$dose==100, 1, 0)*d$recspline

fit1 <- lmer(plasma ~ 1 + fdose* dft + recov30 + recov100 + fnewd + (1|fsubj), data=d)

summary(fit1)


### EXTRACT COEFFICIENTS FROM LMER FIXED EFFECTS OUTPUT FOR
### PLOTS OF LINEAR TRENDS
coef <- summary(fit1)$coef

### CREATE DATA COLUMN CONTAINING MODEL-PREDICTED
### DAY-SPECIFIC MEAN PChE ADJUSTMENTS (FOR LAB TREND)
dayindex <- sort(unique(d$newdate))

dayadj <- NULL
dayadj[1] <- 0
dayadj[2:22] <- coef[8:28]

d$labadj <- dayadj[match(d$newdate, dayindex)]


### CREATE DATA COLUMN CONTAINING SUBJECT-SPECIFIC
### RANDOM EFFECTS FOR PChE ADJUSTMENTS
subranef_df <- ranef(fit1)$fsubj

d$subadj <- subranef_df[match(d$subject, subs),1]

### ADJUST ALL PChE MEASUREMENTS FOR SUBJECT RANDOM EFFECTS
### AND DAY-OF-STUDY LAB VARIATION
d$adjplasma <- d$plasma - d$labadj - d$subadj


###=========================================================
### II. THIS SECTION PLOTS THE ORIGINAL DATA WITH MEAN LINES
###     AND MODEL OVERLAYS
```

```
###============================================================


# day1 CONTAINS THE DAY OF STUDY ON WHICH THE TREATMENT PERIOD
# BEGAN FOR EACH OF THE FOUR DOSE GROUPS; THAT IS, IT IS THE
# DAY CONTAINING THE TIME POINT 24 HOURS AFTER THE FIRST
# ADMINISTRATION OF TREATMENT.
# day0 <- c(7,28,7,14)
day1 <- c(8,29,8,15)

# xdifftrt is the length of the range of dates over which treatment
# period PChE is to be plotted (i.e. from trt day1 to lasttrt)
xdifftrt <- lasttrt - 1


# col, lwd, lty, pch, cex, and smf SPECIFY PLOT PARAMETERS
# FOR COLOR, LINE WIDTH, LINE TYPE, PLOT CHARACTER,
# PLOT CHARACTER SIZE, AND LOWESS SMOOTHING FACTOR (0<smf)

col <- c(1,4,5,6)
lwd <- rep(0.6,4)
lty <- c(1,2,4,5)
pch <- c(1,2,3,4)
cex <- rep(0.5,4)
smf <- 0.3


### COMPUTE MEAN LINE COORDINATES ADJUSTED FOR LAB TREND BY DATE

adjmn0 <- tapply(d[d$dose==0 & d$newdate >= day1[1],]$adjplasma,d[d$dose==0 &
d$newdate >= day1[1],]$newdate,mean)
adjmn14 <- tapply(d[d$dose==14 & d$newdate >= day1[2],]$adjplasma,d[d$dose==14 &
d$newdate >= day1[2],]$newdate,mean)
adjmn30 <- tapply(d[d$dose==30 & d$newdate >= day1[3],]$adjplasma,d[d$dose==30 &
d$newdate >= day1[3],]$newdate,mean)
adjmn100 <- tapply(d[d$dose==100 & d$newdate >= day1[4],]$adjplasma,d[d$dose==100 &
d$newdate >= day1[4],]$newdate,mean)


############################
#### PLOTTING STARTS HERE
############################
```

```
### PLOT CONTROL GROUP PChE DATA, MEAN LINE, AND MODEL OVERLAYS ------
--------------------

par(mar=c(4,4,2,2))
par(ps=10,cex=1,cex.main=1)
plot(d$newdate,d$adjplasma,xlim=c(0,64),ylim=c(1,7),type="n",ylab="Adjusted PChE Activity
(µmol acetate/min/ml)", xlab="Day of Study")

for (i in 1:4){
  lines(d[d$subject==subs0[i] & d$newdate >= day1[1],]$newdate,d[d$subject==subs0[i] &
d$newdate >= day1[1],]$adjplasma,col=1,lty=1,lwd=0.5)
  points(d[d$subject==subs0[i] & d$newdate >= day1[1],]$newdate,d[d$subject==subs0[i] &
d$newdate >= day1[1],]$adjplasma,col=1,cex=cex[i],pch=pch[i])
  lines(d[d$subject==subs0[i] & d$newdate < day1[1],]$newdate,d[d$subject==subs0[i] &
d$newdate < day1[1],]$adjplasma,col=1,cex=cex[i],lty=lty[i])
  points(d[d$subject==subs0[i] & d$newdate < day1[1],]$newdate,d[d$subject==subs0[i] &
d$newdate < day1[1],]$adjplasma,col=1,cex=cex[i],pch=pch[i])
}

endrec0 <- max(d$recspline[d$dose==0])

x1_0 <- day1[1]
x2_0 <- x1_0 + xdifftrt[1]- 7 ## The -7 adjustment corrects error from absence of knot in the
control group


y1_0 <- coef["(Intercept)","Estimate"]+coef["dft","Estimate"]
y2_0 <- y1_0 + coef["dft","Estimate"]*(xdifftrt[1]-7)

## NEXT LINE PLOTS UNADJUSTED MEAN
# lines(lowess(as.numeric(names(means0)),means0,f=smf),col=1,lwd=3)
## NEXT LINE PLOTS LAB-ADJUSTED MEAN
lines(lowess(as.numeric(names(adjmn0)),adjmn0,f=smf),col=1,lwd=3)

### PLOT MODELED LINEAR TREND LINES
lines(c(x1_0,x2_0),c(y1_0,y2_0),lty=1,lwd=1.5,col=4)



### PLOT 14µg GROUP PChE DATA, MEAN LINE, AND MODEL OVERLAYS ----------
-------------

par(mar=c(4,4,2,2))
par(ps=10,cex=1,cex.main=1)
plot(d$newdate,d$adjplasma,xlim=c(0,64),ylim=c(1,7),type="n",ylab="Adjusted PChE Activity
(µmol acetate/min/ml)", xlab="Day of Study")
```

```
for (i in 1:4){
  lines(d[d$subject==subs14[i] & d$newdate >= day1[2],]$newdate,d[d$subject==subs14[i] &
d$newdate >= day1[2],]$adjplasma,col=4,lty=5,lwd=0.5)
  points(d[d$subject==subs14[i] & d$newdate >= day1[2],]$newdate,d[d$subject==subs14[i] &
d$newdate >= day1[2],]$adjplasma,col=4,cex=cex[i],pch=pch[i])
  lines(d[d$subject==subs14[i] & d$newdate < day1[2],]$newdate,d[d$subject==subs14[i] &
d$newdate < day1[2],]$adjplasma,col=4,cex=cex[i],lty=lty[i])
  points(d[d$subject==subs14[i] & d$newdate < day1[2],]$newdate,d[d$subject==subs14[i] &
d$newdate < day1[2],]$adjplasma,col=4,cex=cex[i],pch=pch[i])
}

endrec14 <- max(d$recspline[d$dose==14])

x1_14 <- day1[2]
x2_14 <- x1_14 + xdifftrt[2]
x3_14 <- x2_14 + endrec14

y1_14 <- coef["(Intercept)","Estimate"] + coef["fdose14","Estimate"] +
(coef["dft","Estimate"]+coef["fdose14:dft","Estimate"])
y2_14 <- y1_14 + (coef["dft","Estimate"]+coef["fdose14:dft","Estimate"])*xdifftrt[2]
## y3_14 <- y2_14 +
(coef["dft","Estimate"]+coef["fdose30:dft","Estimate"]+coef["recov30","Estimate"])*endrec14

## OPTION: NEXT LINE PLOTS UNADJUSTED MEAN
# lines(lowess(as.numeric(names(means14)),means14,f=smf),col=4,lwd=3)

## NEXT LINE PLOTS LAB-ADJUSTED MEAN
lines(lowess(as.numeric(names(adjmn14)),adjmn14,f=smf),col=4,lwd=3)

### PLOT MODELED LINEAR TREND LINES
lines(c(x1_14,x2_14),c(y1_14,y2_14),lty=1,lwd=1.5,col=1)
# lines(c(x2_14,x3_14),c(y2_14,y3_14),lty=1,lwd=1.5,col=1)



### PLOT 30µg GROUP PChE DATA, MEAN LINE, AND MODEL OVERLAYS ----------
----------

par(mar=c(4,4,2,2))
par(ps=10,cex=1,cex.main=1)
plot(d$newdate,d$adjplasma,xlim=c(0,64),ylim=c(1,7),type="n",ylab="Adjusted PChE Activity
(µmol acetate/min/ml)", xlab="Day of Study")

for (i in 1:4){
```

```
  lines(d[d$subject==subs30[i] & d$newdate >= day1[3],]$newdate,d[d$subject==subs30[i] &
d$newdate >= day1[3],]$adjplasma,col=3,lty=lty[i],lwd=0.8)
  points(d[d$subject==subs30[i] & d$newdate >= day1[3],]$newdate,d[d$subject==subs30[i] &
d$newdate >= day1[3],]$adjplasma,col=3,cex=cex[i],pch=pch[i])
  lines(d[d$subject==subs30[i] & d$newdate < day1[3],]$newdate,d[d$subject==subs30[i] &
d$newdate < day1[3],]$adjplasma,col=3,cex=cex[i],lty=lty[i])
  points(d[d$subject==subs30[i] & d$newdate < day1[3],]$newdate,d[d$subject==subs30[i] &
d$newdate < day1[3],]$adjplasma,col=3,cex=cex[i],pch=pch[i])
}

endrec30 <- max(d$recspline[d$dose==30])

x1_30 <- day1[3]
x2_30 <- x1_30 + xdifftrt[3]
x3_30 <- x2_30 + endrec30

y1_30 <- coef["(Intercept)","Estimate"] + coef["fdose30","Estimate"] +
(coef["dft","Estimate"]+coef["fdose30:dft","Estimate"])
y2_30 <- y1_30 + (coef["dft","Estimate"]+coef["fdose30:dft","Estimate"])*xdifftrt[3]
y3_30 <- y2_30 +
(coef["dft","Estimate"]+coef["fdose30:dft","Estimate"]+coef["recov30","Estimate"])*endrec30

## NEXT LINE PLOTS LAB-ADJUSTED MEAN
lines(lowess(as.numeric(names(adjmn30)),adjmn30,f=smf),col=3,lwd=2)

### PLOT MODELED LINEAR TREND LINES
lines(c(x1_30,x2_30),c(y1_30,y2_30),lty=1,lwd=1.5,col=1)
lines(c(x2_30,x3_30),c(y2_30,y3_30),lty=1,lwd=1.5,col=1)



### PLOT 100µg GROUP PChE DATA, MEAN LINE, AND MODEL OVERLAYS ----------
----------

par(mar=c(4,4,2,2))
par(ps=10,cex=1,cex.main=1)
plot(d$newdate,d$adjplasma,xlim=c(0,64),ylim=c(1,7),type="n",ylab="Adjusted PChE Activity
(µmol acetate/min/ml)", xlab="Day of Study")

for (i in 1:4){
  lines(d[d$subject==subs100[i] & d$newdate >= day1[4],]$newdate,d[d$subject==subs100[i] &
d$newdate >= day1[4],]$adjplasma,col=6,lty=lty[i],lwd=lwd[i])
  points(d[d$subject==subs100[i] & d$newdate >= day1[4],]$newdate,d[d$subject==subs100[i]
& d$newdate >= day1[4],]$adjplasma,col=6,cex=cex[i],pch=pch[i])
 lines(d[d$subject==subs100[i] & d$newdate < day1[4],]$newdate,d[d$subject==subs100[i] &
d$newdate < day1[4],]$adjplasma,col=6,cex=cex[i],lty=lty[i])
```

```
 points(d[d$subject==subs100[i] & d$newdate < day1[4],]$newdate,d[d$subject==subs100[i] &
d$newdate < day1[4],]$adjplasma,col=6,cex=cex[i],pch=pch[i])
}

endrec100 <- max(d$recspline[d$dose==100])

x1_100 <- day1[4]
x2_100 <- x1_100 + xdifftrt[4]
x3_100 <- x2_100 + endrec100

y1_100 <- coef["(Intercept)","Estimate"] + coef["fdose100","Estimate"] +
(coef["dft","Estimate"] + coef["fdose100:dft","Estimate"])
y2_100 <- y1_100 + (coef["dft","Estimate"] + coef["fdose100:dft","Estimate"])*xdifftrt[4]
y3_100 <- y2_100 + (coef["dft","Estimate"] + coef["fdose100:dft","Estimate"] +
coef["recov100","Estimate"])*endrec100

## NEXT LINE PLOTS UNADJUSTED MEAN
# lines(lowess(as.numeric(names(means100)),means100,f=smf),col=6,lwd=3)
## NEXT LINE PLOTS LAB-ADJUSTED MEAN
lines(lowess(as.numeric(names(adjmn100)),adjmn100,f=smf),col=6,lwd=2)

### PLOT MODELED LINEAR TREND LINES
lines(c(x1_100,x2_100),c(y1_100,y2_100),lty=1,lwd=1.5,col=1)
lines(c(x2_100,x3_100),c(y2_100,y3_100),lty=1,lwd=1.5,col=1)
```

**3b. Confidence intervals for recovery model slope estimates**

```
#########################################################################
##### USES OUTPUT FROM TREATMENT/RECOVERY MODEL 'fit1' #####
#########################################################################

## Access treatment/recovery model coefficient table and variance/covariance matrix
coef <- summary(fit1)$coef
vcov <- vcov(fit1)


## Treatment slope and SE estimates for each dose group

tr0 <- coef["dft","Estimate"]
setr0 <- coef["dft","Std. Error"]

tr14 <- coef["fdose14:dft","Estimate"]
setr14 <- coef["fdose14:dft","Std. Error"]
```

```
tr30 <- tr0 + coef["fdose30:dft","Estimate"]
setr30 <- coef["fdose30:dft","Std. Error"]

tr100 <- tr0 + coef["fdose100:dft","Estimate"]
setr100 <- coef["fdose100:dft","Std. Error"]


## Recovery slope and SE estimates for each dose group

rc30 <- tr30 + coef["recov30","Estimate"]
varrc30 <- vcov["fdose30:dft","fdose30:dft"] + vcov["recov30","recov30"] +
2*vcov["recov30","fdose30:dft"]
serc30 <- sqrt(varrc30)

rc100 <- tr100 + coef["recov100","Estimate"]
varrc100 <- vcov["fdose100:dft","fdose100:dft"] + vcov["recov100","recov100"] +
2*vcov["recov100","fdose100:dft"]
serc100 <- sqrt(varrc100)


## Compute 95% confidence intervals for all treatment slope estimates
citr0 <- tr0+setr0*c(-1,1)*qnorm(0.975)
citr14 <- tr14+setr14*c(-1,1)*qnorm(0.975)
citr30 <- tr30+setr30*c(-1,1)*qnorm(0.975)
citr100 <- tr100+setr100*c(-1,1)*qnorm(0.975)

## Compute 95% confidence intervals for 30µg and 100µg recovery slope estimates
circ30 <- rc30+serc30*c(-1,1)*qnorm(0.975)
circ100 <- rc100+serc100*c(-1,1)*qnorm(0.975)


## Construct and print output table

output_tr <- rbind(c(tr0,citr0),c(tr14,citr14),c(tr30,citr30),c(tr100,citr100))
output_rc <- rbind(rep(0,3),c(rc30,circ30),c(rc100,circ100))

neatoutput <- round(rbind(output_tr, output_rc),3)
print(neatoutput)
```