A FRESH LOOK AT THE RANCH HAND DATA: COMPOSITE BIRTH DEFECTS AND DEVELOPMENTAL DISABILITIES

George J. Knafl and Linda S. Schwartz

Yale School of Nursing, 100 Church Street South, P.O. Box 9740, New Haven, CT 06536-0740

Introduction

The Air Force Health Study has investigated for over 20 years the impact of herbicide exposure on the health, survival, and reproductive outcomes of male Air Force Vietnam War veterans.[1] Reproductive outcomes analyses utilizing serum dioxin, published in 1995, concluded that the data provided little or no support for the theory that paternal exposure to Agent Orange and its dioxin contaminant is associated with adverse reproductive outcomes.[2] Study data released to the public within the last few years were used to conduct a secondary analysis of composite categories of birth defects and developmental disabilities, addressing the impact of paternal dioxin exposure during Vietnam service. This analysis used all available data, encompassing more subjects and births than used in previous reported analyses while considering more possible covariates including ones accounting for temporal inter-sibling dependence. Furthermore, the dose-response relationship was modeled using a novel approach based on a threshold for nonnegligible dioxin exposure and a possible nonlinear effect to such nonnegligible levels as a way to estimate that relationship adaptively in terms of observed serum dioxin however it might have been affected by decay since Vietnam service.

Methods and Materials

Study subjects included personnel responsible for herbicide handling and spraying in Vietnam (Operation Ranch Hand) and a comparison group of other Air Force Vietnam veterans. The reproductive outcomes data consisted of 8,437 pregnancies with 6,923 of these resulting in live births fathered by 2,614 subjects and conceived over the 58 year period from 9/1933 to 8/1991. The conception date was missing for one birth, so it was not used. The analysis addressed the occurrence within live births of nine composite categories of defects and disabilities: defects (any within 12 categories), disabilities (any within 4 categories), defects or disabilities, multiple defects, multiple disabilities, multiple defects or disabilities, major defects, minor defects, and defects of unclassified severity (10.5% of births with defects were not classified as either major or minor). Disabilities were not classified on severity, and so this was not addressed.

Thirty-four variables were considered as possible covariates. Eight of these variables accounted for occurrences of defects/disabilities and live births in prior and concurrent (e.g., twin) conceptions fathered by the same subject (together with associated missing value indicators), thereby accounting for temporal inter-sibling dependence. Linear and quadratic transforms of conception date were included to account for possibly nonlinear background changes over time. Six variables measured aspects of the pregnancy (e.g., single vs. plural, gestation period, birth weight, and sex). Seven variables measured aspects of the mother (e.g., spouse or not, smoked/drank during pregnancy or not, age at birth). The remaining 11 variables measured aspects of the subject (e.g., black/flying/officer or not, age at birth, age at start of tour, tour length, and dioxin exposure). Only the month and year were available for the beginning and end of Vietnam service, so age at the start of the tour and tour length were computed as if the tour lasted for all of any partial months. Four live births were conceived in the month the tour started and were treated as conceived before the start of the tour. The majority of the live births (60.4%) were conceived before the start of the tour.

Serum dioxin levels d were obtained from blood draws taken at various times between 4/1987 and 3/1993, but not for all subjects. The smallest nonzero recorded value was 1.25 ppt; values below the detection limit were set to zero. An initial visualization of the form for the dose-response relationship was obtained by grouping adjacent dioxin exposures, when available, separately for births conceived before/after the start of the tour and plotting percent occurrences for the most inclusive case of any type of defect/disability (Figure 1). Each observation was plotted at its own dioxin value and its group's percent occurrence, and separate nonparametric regression curves were fit to the data for each of the two groups. There are only two points plotted at d=0, one for each conception group, but these correspond to many births, specifically, 129 and 170 births conceived respectively before/after the start of the tour for subjects with dioxin levels below the detection limit. This plot suggested a constant dioxin exposure effect on births conceived before the start of the tour and a threshold effect on births conceived after, with possible nonlinearity after the cutoff controlling the threshold effect. Thus, serum dioxin was included in models using nonlinear transforms adjusted for nonnegligible levels with form $T(d;c,q)=\max(d-c,0)^q$ where c represents a nonnegative cutoff for nonnegligible dioxin levels and q is the nonnegative power controlling the level of nonlinearity. Transformed nonnegligible dioxin exposure T(d;c,q) was considered as a possible covariate to account for possible chance background dependence on dioxin exposure although none was expected. It was also considered interacting with the indicator variable 1_{after} for conception after the start of the tour to account for a Vietnam service dioxin exposure impact. Thus, the Vietnam service dioxin exposure impact represents change from a possibly nonconstant background impact on all births to an adjusted dioxin exposure impact on births conceived after the start of the tour. The indicator for missing dioxin as well as its interaction with 1_{after} were also considered as possible covariates thereby controlling when appropriate for a confounding effect due to missing dioxin and possibly an adjustment to such an effect for births conceived after the start of the tour with missing dioxin.

The basic premise of the analysis was: if there was an impact on birth defects and developmental disabilities due to dioxin exposure during Vietnam service, the odds for composite defects and/or disabilities would increase with dioxin exposure for births conceived after the start of the tour after controlling for confounding effects on all births. The magnitude of the Vietnam service dioxin exposure impact was controlled by the slope parameter for the interaction of transformed nonnegligible dioxin T(d;c,q) and the indicator 1_{after} for conception after the start of the tour. Logistic regression models were generated with the intercept and Vietnam service dioxin exposure impact retained in those models together with slopes for covariates chosen by backward elimination. However, it was also necessary to select appropriate values for the cutoff c and the power q. These quantities were chosen adaptively by minimizing the Akaike information criterion score AIC(c,q) using a grid search (over multiples of 0.1 for c and of 0.01 for q).

Results and Discussion

There is a significant impact associated with transformed nonnegligible Vietnam service dioxin exposure for all composite defect/disability categories except minor defects after controlling for appropriate sets of covariates (Table 1). Thus, there is distinct evidence in these data supporting the theory of an adverse impact of paternal dioxin exposure on birth defects and developmental disabilities. This opposite conclusion to previous analyses seems primarily a consequence of consideration of a dose-response relationship that distinguishes between negligible and nonnegligible exposure with possible nonlinearity and perhaps also due to utilizing all available data and accounting for temporal inter-sibling dependence. When the selected power is small or zero as for many of the cases, the impact is essentially all due to nonnegligible exposure above the threshold determined by the selected cutoff. Similar cutoffs (2.2-2.4 ppt) are obtained in all but two cases. Using a cutoff of 2.3, at the middle of these values, with a power of zero and choosing covariates by backward elimination generates almost the same results as in Table 1. All cases are significant ($p<0.05$) except for major defects ($p=0.15$). Consideration of possible nonlinearity allows the backward elimination to refine the covariate selection so that significance can be detected for major defects while nonsignificance is identified for minor defects. Distinguishing between negligible and nonnegligible exposure is crucial for identifying the existence of a distinct dioxin exposure impact while consideration of nonlinear transforms of exposure levels provides a more precise delineation of the extent of that adverse impact.

There is a multiple comparisons issue underlying Table 1 since nine different tests are conducted. However, all the cases are subsumed by the single any defect/disability case, and so its p-value serves as a measure of the joint significance for all tests combined. Since it is highly significant ($p < 0.01$), the issue of multiple comparisons is not a limitation to the results.

Births conceived before/after the start of the tour tended to be conceived earlier/later in the 58 year conception period and so the identified impact could be partially due to a temporal effect. However, linear and quadratic transforms of conception date were considered as possible covariates and were selected in all but two cases, so that such a possible temporal effect has likely been accounted for. Dioxin exposures during the tour would have only affected births conceived after the start of the tour, but confounding effects (due to covariates) would have affected all births; so including births conceived before the start of the tour provides improved estimates of such effects. They also allow the background dioxin impact to be estimated. However, for all Table 1 categories, this impact was estimated to be constant. When the subset of conceptions after the start of the tour is analyzed using Table 1 cutoffs and powers, nearly the same results are produced. All cases are significant ($p<0.05$) except for multiple defects ($p=0.10$). While the significant cases are a little different, a distinct dioxin exposure impact is still identified using only births conceived after the start of the tour. This also indicates that the significant effects identified in Table 1 are not an artefact of distinguishing between conceptions before/after the start of the tour rather than a consequence of dioxin exposure. Another way to address this issue is to replace the Vietnam service dioxin exposure impact by a Vietnam service impact based entirely on conception before/after the tour without considering dioxin levels (measured or missing) for births conceived after the start of the tour. Only one of the Table 1 cases is significant (unclassified defects, $p=0.02$), another is marginally significant (disabilities, $p=0.07$), and the rest are nonsignificant ($p>0.10$) indicating that the Table 1 results are related to dioxin exposure after the start of the tour not to relative timing of conception.

Births conceived for the same father are dependent. An autoregressive logit approach was used to control for this dependence by accounting for defect, disability, parity, and gravida information for previous conceptions and also for dependence within plural births with the same conception date. There was a large percentage (20.9%) of births with missing dioxin and other variables, birth weight (4.1%) and mother's age at birth (1.5%), had missing values as well. However, each of these was coded into two variables: the variable with its missing values reset to zero to use its value whenever measured (formally, the interaction between the completely measured variable and the indicator for when it is actually measured) and an indicator of missing or not to account for cases when not measured. Unknown values for classifications were handled similarly. These adjustments allow analyses to be based on all available births while utilizing all available variables even if incompletely measured. However, whenever models include variables adjusted for missing values, those covariates apply only to observations with measured values. When the subset of births for fathers with measured dioxin levels are analyzed with Table 1 cutoffs and powers, the minor defects case stays nonsignificant, three cases become nonsignificant (multiple defects, $p=0.06$; multiple disabilities, $p=0.08$; major defects, $p=0.13$), and the remaining five cases remain significant ($p<0.05$). More cases might possibly become significant if cutoffs and powers are also adjusted. On the other hand, analyzing the complete data might have provided increased power for detecting significant effects. The complete-data analysis identified significant effects due to missing dioxin in seven cases, and so not accounting for these births might have had an effect on the dioxin-subset analysis. In any case, the three primary cases of any defect, any disability, and any defect/disability remain significant for the dioxin-subset analysis supporting the existence of a distinct Vietnam service dioxin exposure impact but with its full extent somewhat masked by not considering all available births.

Variable selection procedures can be problematic. Large variable sets can include variables that are nearly constant or nearly equal to other variables for most occurrences of a defect/disability type. Backward elimination sometimes generated warning messages about such situations at initial stages, but the warning was resolved before the final stage in all cases. Much different results can sometimes be produced by alternative variable selection procedures. Backward elimination was used to select covariates since it tends to identify more significant effects than forward or stepwise selection. Slope parameters for all covariates of Table 1 models were significantly nonzero ($p<0.05$). When stepwise selection is used instead with Table 1 cutoffs and powers, the same cases are significant, but somewhat different covariate sets are generated. This near consistency in variable selection results provides an indication that appropriate covariates were used to generate Table 1 results. On the other hand, none of the AIC scores for the stepwise case are better than for the backward elimination case, indicating that backward elimination provided a somewhat more thorough covariate identification. Somewhat different results might be produced if the intercept was also considered for elimination. However, the intercept parameter was significant for all but one category (unclassified severity), so it is likely that the results would be affected at most slightly by consideration of zero intercept models.

Serum dioxin was measured long after subjects' Vietnam service, and so observed values very likely reflect distinct amounts of decay. This decay is also likely to have affected the dose-response relationship. A nonlinear model for this relationship was formulated and its parameters were selected adaptively in order to address this issue in terms of observed decayed values without extrapolating them back to the tour. A threshold effect was hypothesized with nonlinear impact to nonnegligible exposure that varied with defect/disability category. The selected values for cutoffs controlling this threshold are most likely smaller due to decay than if exposure had been measured just after the tour. It seems likely though that most initial nonnegligible exposures (at a higher cutoff than estimated) would not have decayed to below estimated nonnegligible levels given dioxin's known persistence in the body. However, if decay has occurred to such a substantial extent that initial nonnegligible exposures are quite likely to have become negligible, the actual dioxin exposure impact would be stronger than the impact identified in Table 1. Dioxin measurements occurred over a six year period with delay between the end of the tour and measurement varying from 9.75 years to 30.4 years. Since these delays are all quite long, the associated decays in dioxin exposure might all be substantial enough to have about the same effects due to delay. On the other hand, the largest delay is more than three times the shortest, and so there may be differential decay effects. To assess this, we considered the most inclusive case of any defect/disability. The Table 1 model for this case depends only on untransformed nonnegligible exposure with cutoff 2.3 ppt. Suppose instead that the cutoff for the threshold effect is given by $2.3/u^{(}$ where u is the delay time divided by its smallest value 9.75 and (controls the level of nonlinear impact to normalized delay. That is, starting at the same exposure level, a larger delay should produce more exposure decay and hence require a lower cutoff (with (>0) to identify the original exposure as nonnegligible. When AIC scores are compared for different choices of (around zero for the any defect/disability case, the best score is attained at (=0. This indicates that a constant threshold for nonnegligible exposure without adjusting for delay in exposure measurement is a reasonable simplification.

The amount of decay should increase the further conception is from Vietnam service, but almost always (only 21 conceptions occurred after measurement) not as much as by the time of measurement, and so births conceived after the start of the tour for subjects with nonnegligble exposure would have been conceived when that exposure was also nonnegligible but with larger levels than as measured. The selected powers controlling the nonlinearity have most likely also been affected by decay in some complex way. On the other hand, exposures might have sometimes occurred outside of the tour. A background dioxin exposure impact was not identified for any of the Table 1 cases. Thus, it seems unlikely that dioxin exposures prior to the tour, if any, occurred at tangible enough levels to have contributed to the identified significantly adverse dioxin impact. Tangible exposures might have occurred though after the tour, but the original study did not attempt to address this issue (2, p. 18) which seems the most serious limitation.

Acknowledgements

Dr. Knafl's research was supported in part by Vietnam Veterans Assistance Fund Grant 01-10.

References

1. http://www.brooks.af.mil/AFRL/HED/hedb/afhs/afhs.html.
2. Wolfe W.H., Michalek J.E., Miner J.C., Rahe A.J., Moore C.A., Needham L.L., Patterson D.G., Jr. (1995) Epidemiology. 6, 17-22.

Figure 1

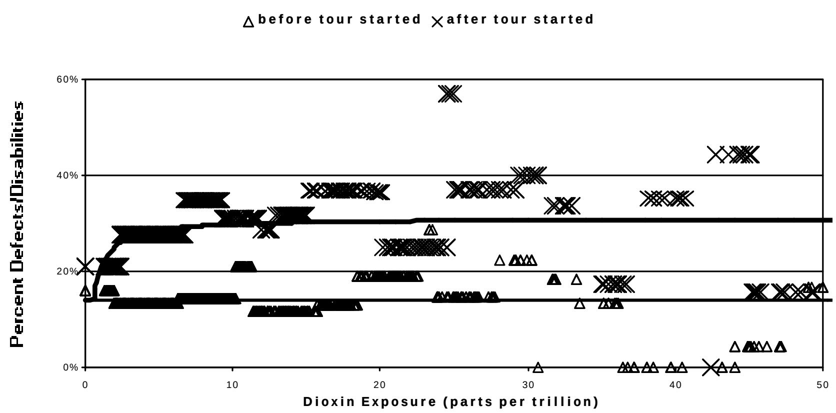

Table 1: Impact Associated with Transformed NonNegligible Vietnam Service Dioxin Exposure on Composite Categories of Birth Defects and Developmental Disabilities

category	c	q	AIC	p	OR	95% CI
defects						
any type	2.3	0.03	5363.371	0.03	1.25	1.02-1.52
multiple	1.9	0.11	1416.869	0.04	1.53	1.01-2.32
defect severity						
major	2.2	0.20	2367.977	0.02	1.26	1.03-1.53
minor	2.2	0.01	3827.003	0.27	1.15	0.90-1.48
unclassified	1.7	0.17	1024.494	<0.01	1.68	1.22-2.29
disabilities						
any type	2.3	0.00	2730.902	0.02	1.44	1.06-1.95
multiple	2.4	0.00	791.088	0.01	2.23	1.23-4.06
either						
any type	2.3	0.00	6087.381	<0.01	1.37	1.14-1.65
multiple	2.3	0.08	2283.354	0.02	1.45	1.06-1.97

c: cutoff in ppt; q: power; AIC: Akaike information criterion; p: p-value; OR: odds ratio for unit increase in transformed nonnegligible dioxin exposure $T(d;c,p)=\max(d-c,0)^q$ at dioxin exposure level d for births conceived after the start of the tour; CI: confidence interval for OR