Investigating the Relationship Between Water Exposures and Adverse Birth Outcomes

Water exposure data was collected in a particular site in NC and a site in TX as part of study to evaluate the relationship between certain disinfection by-product (DBP) compounds in drinking water and adverse birth outcomes. These compounds are created when chlorine reacts with natural organic matter during water disinfection. The exposure variables of interest include total trihalomethanes (TTHMs), haloacetic acids (HAAs), and total organic halides (TOX). It is thought that the amount and type of these compounds ingested during the second trimester of pregnancy could have an impact on birth outcomes, such as small for gestational age and preterm birth.

A previous study found some of these water exposures to be significant when looking at small for gestational age and preterm birth. That study recruited pregnant women and women planning to become pregnant. Large amounts of exposure data was collected by ongoing surveys throughout pregnancy on these women. The information obtained includes but is not limited to the amount of water ingested, length and temperature of showers and baths, frequency of boiling water, along with other non-water related covariates of interest. The sample sizes were limited in this study. During the study, water samples were collected weekly in the respective sites. This water information was used to create weekly measures of the exposures for the subjects in the study. Those weekly measures were then used to create trimester average water exposure levels. Specifically, the second trimester exposure levels were considered the most important in modeling small for gestation age.

Because of the findings from the study mentioned above, it was of interest to look into this water exposure further. Water exposures collected in that study were used for the analysis to be presented here. The water data was collected from 2000 to 2004 and the exposures were relevant for any individuals living in those two areas during water collection. All births occurring during water collection in the North Carolina and Texas sites were used in the new study. The NC site was chosen because of its moderate levels of chlorinated DBPs and low levels of brominated DBPs. The TX site was chosen because of its moderate levels of brominated DBPs and low levels of chlorinated DBPs. Here, moderate means that the levels approach but do not exceed the levels established by the EPA. It would be much more informative to have very detailed water exposure information on subjects such as that used in the previous study, but collecting that type of information would have been very costly on a large scale.

The exposure variables used in the analysis include categorized and continuous versions of TTHM, HAA5, and TOX. The categorized levels are given in table 1, below. The covariates available here include variables which are recorded on birth certificates, as this is public record and accessible. Logistic regression was used to explore the relationship between small for gestation age with exposures in the drinking water and other covariates of interest. The covariates which are available include maternal age, maternal race, maternal education, maternal alcohol consumption, maternal tobacco use, marital status, and number of mother’s previous births.

|  |  |  |  |
| --- | --- | --- | --- |
| Table 1: Hoffman et al. Categorized | | | |
|  | Exposure Groups\* | | |
| Groups | TTHM | HAA5 | TOX |
| 1 | (0, 4.6] | (0, 0.9] | (0, 22.4] |
| 2 | (4.6, 55] | (0.9, 22] | (22.4, 169.6] |
| 3 | (55, 66.3] | (22, 31.5] | (169.6, 177.7] |
| 4 | (66.3, 74.8] | (31.5, 40.4] | (177.7, 192.6] |
| 5 | > 74.8 | > 40.4 | >192.6 |
| \* Intervals augmented to eliminate gaps. | | | |

The program R was used to analyze the data. The glm procedure was used to implement a logistic regression model of the relationship between the covariates of interest and water exposures and small for gestational age. One data problem was a large number of birth records missing parity information for the TX site. Parity was needed to obtain a value of small for gestational age. SAS was used to impute values of parity for these records. Though it can cause bias, these values were rounded so that small for gestational age could be calculated.

All exposures were kept in the model, regardless of their significance. The model was reduced by eliminating covariates (one at a time) who had the highest p-value, for p-values greater than 0.05. The significance of the water exposures are given in table 2, below. We can see that the only exposure which is significant in modeling small for gestational age for the North Carolina site is the categorical TTHM exposure. The only water exposure which is significant in modeling small for gestational age in the Texas site is the continuous TOX exposure. The subset of covariates significant in modeling small for gestational age was different for the two sites. Though the specific p-values were different in each model, the overall idea of the level of significance for the covariates by site is given in table 3. Table 4 gives the parameter estimates for the two marginally significant exposure variables. From the table, we can see that in the NC site, when comparing the third group to the second (there were no subjects with exposure values in the first group) the estimated odds of small for gestational age increases. When comparing the two higher groups to group 2, however, the estimated odds of small for gestational age decreases. For the TX site, as the amount of TX exposure increases the odds of small for gestational age increases.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Table 2: Significance of Exposures in modeling Small for Gestational Age (p-values) | | | | | | | | |
|  | THM4 Exposure Variable | |  | HAA5 Exposure Variable | |  | TOX Exposure Variable | |
|  | NC | TX |  | NC | TX |  | NC | TX |
| Categorical | 0.0948 | 0.5509 |  | 0.5643 | 0.9718 |  | 0.7449 | 0.5474 |
| Continuous | 0.9645 | 0.2601 |  | 0.7631 | 0.4581 |  | 0.8828 | 0.0725 |

|  |  |  |
| --- | --- | --- |
| Table 3: Significance of Covariates in the Model (p-values) | | |
|  | NC site | TX site |
| Maternal Age (categorized) | <0.001 | - |
| Maternal Race | <0.0001 | <0.01 |
| Marital Status | <0.01 | - |
| Maternal Education Level (categorized) | <0.0001 | <0.05 |
| Maternal Tobacco Use | <0.0001 | <0.0001 |
| Parity | <0.01 | - |

|  |  |  |  |
| --- | --- | --- | --- |
| Table 4: Exposure Model Estimates | | | |
|  |  | Estimate | SE |
| NC THM4 | |  |  |
|  | 3 | 0.1089 | 0.0604 |
|  | 4 | -0.0769 | 0.0741 |
|  | 5 | -0.0139 | 0.0623 |
| TX TOX | |  |  |
|  | cont. | 0.056 | 0.0311 |