National Human Exposure Assessment Survey (NHEXAS)

Arizona Study

Quality Systems and Implementation Plan for Human Exposure Assessment

The University of Arizona
Tucson, Arizona 85721

Cooperative Agreement CR 821560

Standard Operating Procedure   **SOP-IIT-A-4.0**

Title: Treatment of Censored Data

Source: The University of Arizona

U.S. Environmental Protection Agency
Office of Research and Development
Human Exposure & Atmospheric Sciences Division
Human Exposure Research Branch

STANDARD OPERATING PROCEDURE
FOR
TREATMENT OF CENSORED DATA

This Standard Operating Procedure (SOP) uses data that have been properly coded and certified with appropriate QA/QC procedures by the University of Arizona NHEXAS team.

Objective

Censor data which are below detection limit.

Approaches

Censored data are pesticide concentrations that have been qualified as Below Detection Limit (BDL) values. Four different approaches will be followed to deal with censored pesticide concentration data that for certain cases constitute a very large segment of the data collected.

Approach 1: Substitution with zero values

Assign the value zero to all censored data. This is a very liberal approach in the sense that it minimizes the population exposure.

Approach 2: Substitution with the detection limit value

Assign the value of the limit of detection to all censored data. In the case of multiple detection limits, assign the highest limit of detection to all censored data. If the minimum detectable value is lower than the maximum non-detectable values, use the minimum detectable value as the detection limit for all censored data. This is a conservative approach that maximizes the population exposure to pesticides.

Approach 3: Robust Method

Use the Crystal Ball software to fit the non-censored data with the best fitting function, and assume that all data, above and below the limit of detection follow the same distribution. For purposes of this study, the goodness of fit will be tested by chi-squared test, Kolmogorov-Smirnov tests, and Anderson-Darling test. At least one of these tests must consider the fit acceptable. This criterion, however, is relaxed because of the second

criterion: the first criterion estimates the best fit for data above the detection limit, but the second one requires that the percentage of censored data be considered when the final selection is made. The two criteria combined lead usually to selection of a lognormal distribution which may not always be the best fit but satisfies both criteria.

Once the best-fit distribution is obtained, Monte Carlo techniques will be used to select values from the censored portion of the distribution and randomly associate them with subjects in residences with censored pesticide levels. This step creates "fill-in" values, based on extrapolation of the fitted distribution, for samples which are below the detection limit. The observed data above the detection limit and the "fill-in" values will be used when estimating summary statistics.

Approach 4: Distributional Method

This approach uses the same method as in Approach 3 to determine the best-fit distribution. Once the best-fit distribution is obtained, its characteristics will be used when estimating summary statistics. The fitted distributions will be used in the estimation of exposures using the probabilistic approach.