

Pesticide Risk Mitigation Engine

Policy on Missing Data Points

Goal: To ensure the least degree of bias resulting from missing data.

General policies:

1. We will not replace missing data with a worst case value. Data points may be missing intentionally because the compound in question is of minimal concern. For example, insecticides are not generally thought to be toxic to plants and may therefore not have been tested for their phytotoxicity; i.e. an algal toxicity test may be missing.
2. Missing data will be replaced as follows (in priority order):
 - a. By data approximated from a robust model derived from other products including Quantitative Structure Activity Relationships (QSAR) approaches (Note: QSAR models are mathematical relationships that link chemical structure with physical or toxicological activity in a quantitative fashion for a large number of known chemicals.)
 - b. By data from a closely related compound/species
 - c. By data averaged among a group of moderately close compounds (based on broad chemical groupings)
3. When different approaches result in dramatically different rankings, fields and results may be left blank, warning the user of the missing data.
4. When data are approximated, this use will be clear and transparent to the user.