Standardized Data Evaluation Can Enhance Quality and Throughput

Christopher Bevan, Ph.D., D.A.B.T.
CJB Consulting LLC

Introduction

Chemical risk assessments are intended to provide sufficient information to evaluate potential hazard associated with relevant exposures. They should:

- Use best available scientific information as the basis for conclusions
  - Most reliable and scientifically support information
  - High standard for quality and completeness of data
- Offer relevant and useful information for sound decision-making by regulators
- Convey appropriate risk information to public
- Identify uncertainties and limitations in assessment

Introduction (cont'd)

- Risk assessor must consider many study characteristics when evaluating studies
  - Study design and methodologies
  - Study quality
- Standardized procedure for data evaluation is necessary
  - Ensures consistency and transparency
  - Credibility, objectiveness, and scientifically supported assessment

Data Quality Evaluation for Hazard/Risk Characterization

White paper developed on “best practices”

- Reviewed procedures with formal and documented processes
- Focused on *In vivo* and *in vitro* mammalian toxicity studies only
- Identified critical aspects that should be considered in the data quality evaluation
  - Method validity
  - Reproducibility
  - Study reliability
  - Appropriateness of study for risk assessment
- Sponsored by American Chemistry Council (ACC)’s Center for Advancing Risk Assessment Science and Policy (ARASP)

Data Quality Evaluation for Hazard/Risk Characterization (cont'd)

Overarch Takeaway from white-paper

- Study data should meet minimum data quality requirements to be considered
- Expert judgment required for evaluation of data quality/relevance from studies using non-standard test guideline methods
  - Suitability of data for hazard identification and for quantification of risk
  - Expectation that objective, transparent, well-documented justification of data will be accepted by independent peer reviewers
- Data quality evaluation methods are currently available and could be readily implemented

Klimisch Code System

- Klimisch scoring system is a structured approach used for reviewing data for existing chemicals and currently in use:
  - U.S. EPA and OECD HPV programs
  - EU REACH registrations
- Klimisch *et al.* (1997) defined three aspects for evaluating data quality for hazard/risk assessments
  - Reliability
    - Inherent quality of a test report/publication relating to standard methodology
    - Description lf experimental procedures/results to give clarity and plausibility of the findings
  - Relevance
    - Extent to which data and/or tests are appropriate for a particular hazard identification/risk characterization
  - Adequacy
    - Usefulness of data for risk assessment purposes
    - Data weighted based on reliability and relevance

Klimisch *et al.* (1997) Regulatory Toxicology and Pharmacology 25: 1-5

Klimisch Code System (cont'd)

- Scoring system (with justification phrases) developed to assess reliability of toxicology studies
  - 1 (reliable without restriction)
    - Guideline study (OECD, EPA, etc.)
    - Comparable to guideline study
    - Test procedure according the national standards
    - Test procedure in accordance with generally accepted scientific standards and described in sufficient detail
  - 2 (reliable with restriction)
    - Acceptable, well-documented publication/study report which meets scientific principles
    - Basic data given; comparable to guidelines/standards
    - Comparable to guideline study with acceptable restriction
  - 3 (not reliable)
    - Method not validated
    - Documentation insufficient for assessment
    - Does not meet important criteria of today standard methods
    - Relevant methodological deficiencies
    - Unsuitable test system
  - 4 (not assignable)
    - Only short abstract available
    - Only secondary literature

Klimisch Code System (cont'd)

- Greatest weight attached to studies that are most relevant and reliable when multiple studies are available
- Studies with Kl. scores 3 or 4 not necessarily excluded from hazard/risk assessment
  - Used as supporting evidence
  - Weight-of-evidence approach
  - Sound scientific judgment needed
- Limitations:
  - Not structured to evaluate studies that do not have internationally accepted guidelines
  - Determination of key/critical studies when multiple studies are available which are of comparable reliability
  - Question of adequacy or usefulness of data for hazard/risk assessment only partially addressed

ECVAM ToxRTool

- ToxRTool (Toxicological data Reliability Assessment Tool) free software developed by ECVAM (http://ecvam.jrc.it)
  - Provides criteria and guidance for reliability evaluations for *in vivo* and *in vitro* toxicological data
  - Excel spreadsheet
  - Five evaluation criteria groupings:
    - Test substance identification
    - Test system characterization
    - Study design description
    - Study results documentation
    - Plausibility of study design and data

[1]Schneider *et al.* (2009) Toxicology Letters 189: 138-144.

ECVAM ToxRTool (cont'd)

- For each criteria within a group, "1" is assigned if study meets that criteria, or "0" if it is not met.
- Total number of points establishes a reliability category (the same as the Klimisch codes)
- ToxRTool can be particularly useful for *in vitro* studies
  - List of elements can be developed to help risk assessor design a set of data evaluation procedures for determining the quality and reliability of an *in vitro* study

Key Recommendations

- A systematic approach with objective criteria will be needed to address different types of study data (*e.g.*, *in vitro*, *in vivo*, epidemiological, omics)
  - Klimisch scoring system with expanded list of narratives for studies with internationally accepted guidelines
  - ToxRTool approach for studies that do not have internationally accepted guidelines
    - Such studies be conducted according to scientific acceptable methodology with sufficient documented methods and data
    - Formal quality assurance process or standard provides greater confidence in data quality
- Study quality should be reviewed and considered prior to conducting the weight of evidence analysis

Future Challenges

- Identify standardized evaluation criteria for:
  - Epidemiological studies
  - Nontraditional toxicity tests and toxicity prediction tools
  - Institutionalization and acceptability of standardized criteria to evaluate and weigh studies for appropriateness of use in hazard/risk characterizations