

# **Understanding how Data Quality impacts Decision Quality**

Kara Morgan

Research Triangle Institute

Washington, DC

# Outline

- **Overview**
- **The Problem:** Characteristics of Environmental Decision Making
- **A step in the right direction:**  
Data Quality at EPA
- **How it can be improved:** Planned research
- **Conclusions**

# Overview

- Types of environmental decision making
  - “Statistical assessments”
    - Should we clean up this site?
    - Has regulatory level been exceeded?
  - Policy decision making
    - Should this contaminant be regulated in drinking water?
    - Is there a risk from this contaminant?

# Characteristics of Environmental Decision Making

- “Data Poor”
- Large Uncertainties
  - Information-based
  - Methodological
- High stakes
  - Health
  - Ecological
  - Economic

# Criticisms/Impediments to Good Environmental Decision Making

- Regulatory agency with advocacy role
- Time and budget constraints
- Large scientific uncertainties
- Many stakeholders are involved
- Policy makers lack an understanding of how science can contribute
- Call for more “science-based decision making”

What does science-based  
decision making mean?

# Science-based Decision Making

## (1) Input to the Decision

- Empirical findings that provide insight into the likelihood, desirability, and effectiveness of the available alternatives
- Characterizes uncertainties revealed through research
- A thorough and unbiased review to obtain relevant scientific information

# Science-based Decision Making

## (2) Process

- Uses some type of analysis that considers underlying causative mechanisms and processes
- Includes a structure for incorporating scientific information along with other input
- Based on rigorous application of decision theory

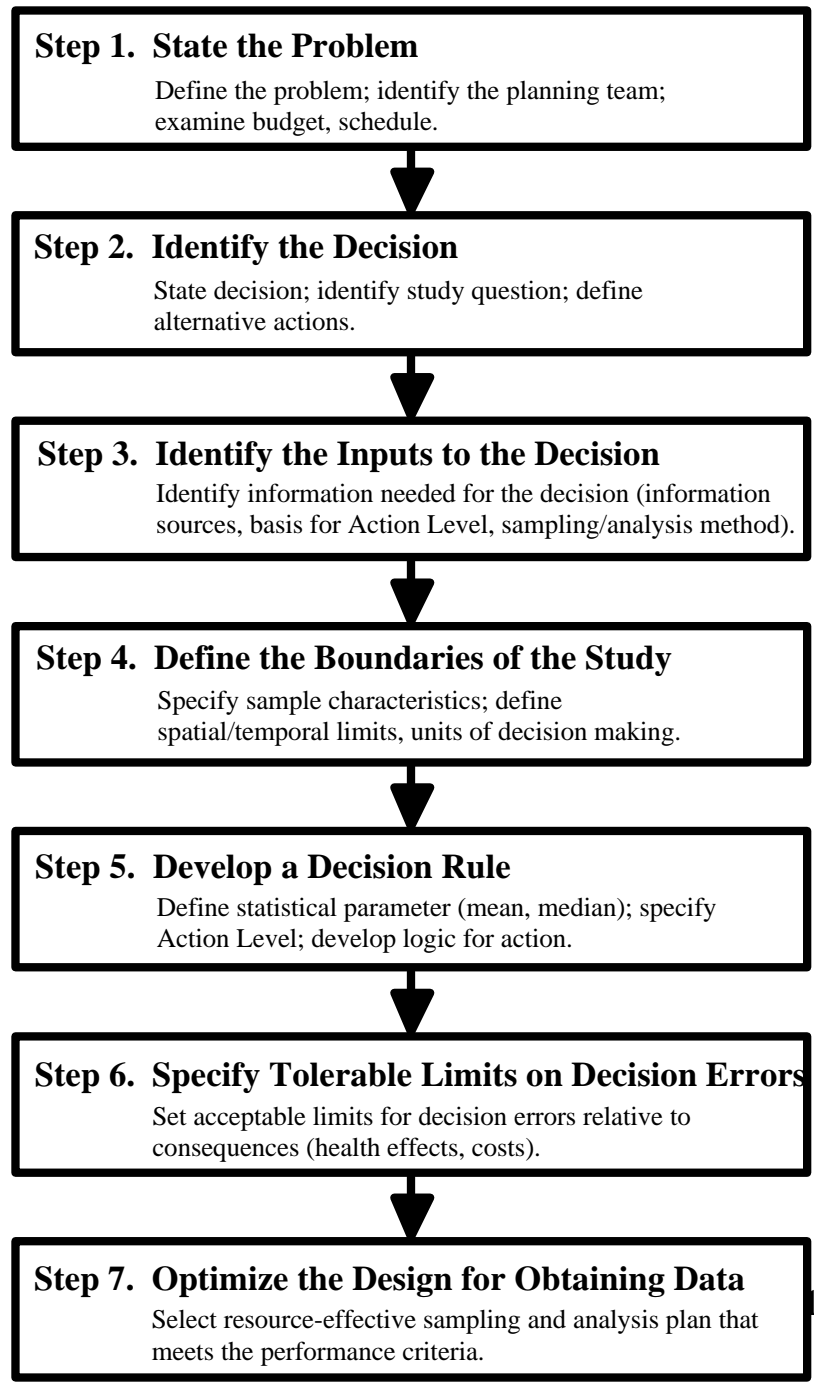
# Good scientific results for decision making are...

- (a) Produced by scientific studies with controls and proper methodologies
- (b) Published in the peer-reviewed literature
- (c) Provide a good estimate of what you need to know for your decision
- (d) All of the above

# EPA's Quality Assurance Program

- Formed by EPA Order 5360.1 in 1984
- Historical documents since 1973
- For the purpose of improving the quality of the data EPA generates for decision making
- Develop guidance & training
- Now in the Office of Environmental Information

# EPA's Data Quality Objectives Process



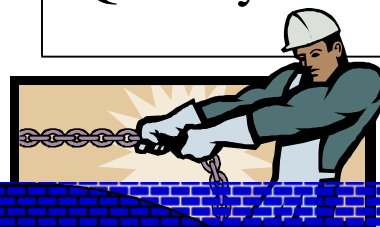
# Data Quality Indicators

- Precision
- Accuracy
- Representativeness
- Completeness
- Comparability
- Selectivity
- Sensitivity

# Challenges

- Negative connotations of “QA”
- Misunderstanding of what is important by those doing data collection
  - QA people are usually chemists
  - Planning, sampling design overlooked
- Under-informed decision makers
  - how to evaluate if data is “good enough”

Building the bridge:  
Decision Analysts  
Statisticians  
Quality assurance professionals



Data Generators:  
Chemists  
Engineers  
Biologists

Decision Makers:  
Project managers  
Policy makers  
Stakeholders

# Goal: Better understanding of how data quality affects decision quality

- On part of data collection people
  - increased understanding will lead to improved quality
- On part of decision makers
  - methods for explicitly incorporating uncertainty based on data quality

# Mental modeling research

- People use tacit webs of belief called “mental models” to make sense of information they use
- Expert model: statisticians, decision analysts
- Lay model: one for data collectors AND one for decision makers
- Compare expert model to lay model

# Develop Training Materials

- To present information from the expert model
  - That reinforces the correct concepts
  - That adds information to fill knowledge gaps
  - That presents information to correct misconceptions
- Information from data collectors and decision makers might also improve the “expert” model.

# Conclusions

- Improving the “scientific-data” part of the process requires social science research
- Improved input to the process
- Improve decision making