

WAC 365-195-905

Criteria for determining which information is the "best available science."

(1) This section provides assessment criteria to assist counties and cities in determining whether information obtained during development of critical areas policies and regulations constitutes the "best available science."

(2) Counties and cities may use information that local, state or federal natural resource agencies have determined represents the best available science consistent with criteria set out in WAC **365-195-900** through **365-195-925**. The department will make available a list of resources that state agencies have identified as meeting the criteria for best available science pursuant to this chapter. Such information should be reviewed for local applicability.

(3) The responsibility for including the best available science in the development and implementation of critical areas policies or regulations rests with the legislative authority of the county or city. However, when feasible, counties and cities should consult with a qualified scientific expert or team of qualified scientific experts to identify scientific information, determine the best available science, and assess its applicability to the relevant critical areas. The scientific expert or experts may rely on their professional judgment based on experience and training, but they should use the criteria set out in WAC **365-195-900** through **365-195-925** and any technical guidance provided by the department. Use of these criteria also should guide counties and cities that lack the assistance of a qualified expert or experts, but these criteria are not intended to be a substitute for an assessment and recommendation by a qualified scientific expert or team of experts.

(4) Whether a person is a qualified scientific expert with expertise appropriate to the relevant critical areas is determined by the person's professional credentials and/or certification, any advanced degrees earned in the pertinent scientific discipline from a recognized university, the number of years of experience in the pertinent scientific discipline, recognized leadership in the discipline of interest, formal training in the specific area of expertise, and field and/or laboratory experience with evidence of the ability to produce peer-reviewed publications or other professional literature. No one factor is determinative in deciding whether a person is a qualified scientific expert. Where pertinent scientific information implicates multiple scientific disciplines, counties and cities are encouraged to consult a team of qualified scientific experts representing the various disciplines to ensure the identification and inclusion of the best available science.

(5) Scientific information can be produced only through a valid scientific process. To ensure that the best available science is being included, a county or city should consider the following:

(a) **Characteristics of a valid scientific process.** In the context of critical areas protection, a valid scientific process is one that produces reliable information useful in understanding the consequences of a local government's regulatory decisions and in developing critical areas policies and development regulations that will be effective in protecting the functions and values of critical areas. To determine whether information received during the public participation process is reliable scientific information, a county or city should determine whether the source of the information displays the characteristics of a valid scientific process. The characteristics generally to be expected in a valid scientific process are as follows:

1. **Peer review.** The information has been critically reviewed by other persons who are qualified scientific experts in that scientific discipline. The criticism of the peer reviewers has been addressed by the proponents of the information. Publication in a refereed scientific journal usually indicates that the information has been appropriately peer-reviewed.

2. **Methods.** The methods that were used to obtain the information are clearly stated and able to be replicated. The methods are standardized in the pertinent scientific discipline or, if not, the methods have been appropriately peer-reviewed to assure their reliability and validity.

3. **Logical conclusions and reasonable inferences.** The conclusions presented are based on reasonable assumptions supported by other studies and consistent with the general theory underlying the assumptions. The conclusions are logically and reasonably derived from the assumptions and supported by the data presented. Any gaps in information and inconsistencies with other pertinent scientific information are adequately explained.

4. **Quantitative analysis.** The data have been analyzed using appropriate statistical or quantitative methods.

5. **Context.** The information is placed in proper context. The assumptions, analytical techniques, data, and conclusions are appropriately framed with respect to the prevailing body of pertinent scientific knowledge.

6. **References.** The assumptions, analytical techniques, and conclusions are well referenced with citations to relevant, credible literature and other pertinent existing information.

(b) **Common sources of scientific information.** Some sources of information routinely exhibit all or some of the characteristics listed in (a) of this subsection. Information derived from one of the following sources may be considered scientific information if the source possesses the characteristics in Table 1. A county or city may consider information to be scientifically valid if the source possesses the characteristics listed in (a) of this subsection. The information found in Table 1 provides a general indication of the characteristics of a valid scientific process typically associated with common sources of scientific information.

Table 1	CHARACTERISTICS					
	Peer review	Methods	Logical conclusions & reasonable inferences	Quantitative analysis	Context	References
SOURCES OF SCIENTIFIC INFORMATION						
A. Research. Research data collected and analyzed as part of a controlled experiment (or other appropriate methodology) to test a specific hypothesis.	X	X	X	X	X	X
B. Monitoring. Monitoring data collected periodically over time to determine a resource trend or evaluate a management program.		X	X	Y	X	X
C. Inventory. Inventory data collected from an entire population or population segment						

(e.g., individuals in a plant or animal species) or an entire ecosystem or ecosystem segment (e.g., the species in a particular wetland).		X	X	Y	X	X
D. Survey. Survey data collected from a statistical sample from a population or ecosystem.		X	X	Y	X	X
E. Modeling. Mathematical or symbolic simulation or representation of a natural system. Models generally are used to understand and explain occurrences that cannot be directly observed.	X	X	X	X	X	X
F. Assessment. Inspection and evaluation of site-specific information by a qualified scientific expert. An assessment may or may not involve collection of new data.		X	X		X	X
G. Synthesis. A comprehensive review and explanation of pertinent literature and other relevant existing knowledge by a qualified scientific expert.	X	X	X		X	X
H. Expert Opinion. Statement of a qualified scientific expert based on his or her best professional judgment and experience in the pertinent scientific discipline. The opinion may or may not be based on site-specific information.			X		X	X

X = characteristic must be present for information derived to be considered scientifically valid and reliable

Y = presence of characteristic strengthens scientific validity and reliability of information derived, but is

not essential to ensure scientific validity and reliability

(c) **Common sources of nonscientific information.** Many sources of information usually do not produce scientific information because they do not exhibit the necessary characteristics for scientific validity and reliability. Information from these sources may provide valuable information to supplement scientific information, but it is not an adequate substitute for scientific information. Nonscientific information should not be used as a substitute for valid and available scientific information. Common sources of nonscientific information include the following:

(i) Anecdotal information. One or more observations which are not part of an organized scientific effort (for example, "I saw a grizzly bear in that area while I was hiking").

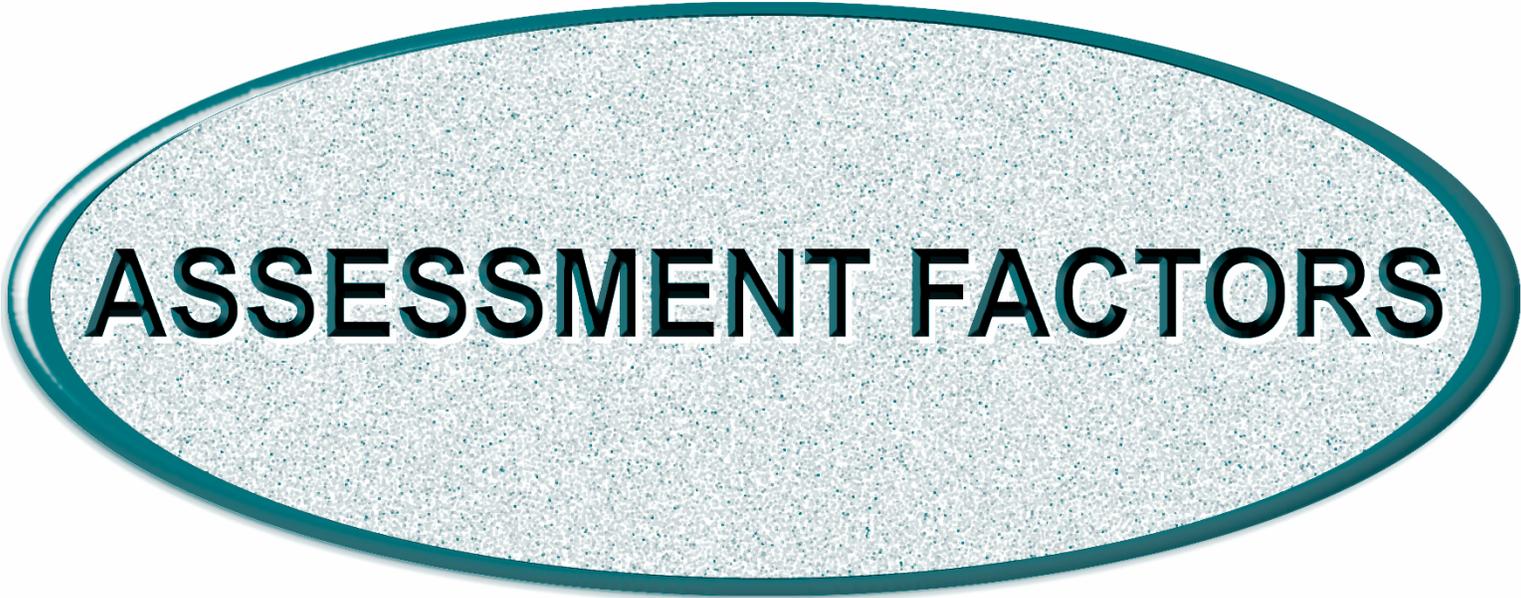
(ii) Nonexpert opinion. Opinion of a person who is not a qualified scientific expert in a pertinent scientific discipline (for example, "I do not believe there are grizzly bears in that area").

(iii) Hearsay. Information repeated from communication with others (for example, "At a lecture last week, Dr. Smith said there were no grizzly bears in that area").

(6) Counties and cities are encouraged to monitor and evaluate their efforts in critical areas protection and incorporate new scientific information, as it becomes available.

[Statutory Authority: RCW **36.70A.190** (4)(b). WSR 00-16-064, § 365-195-905, filed 7/27/00, effective 8/27/00.]

Science Policy Council



ASSESSMENT FACTORS





EPA 100/B-03/001
June 2003

U.S. Environmental Protection Agency

A Summary of General Assessment Factors for Evaluating the Quality of Scientific and Technical Information

**Prepared for the U.S. Environmental Protection Agency
by members of the Assessment Factors Workgroup, a group of the
EPA's Science Policy Council**

**Science Policy Council
U.S. Environmental Protection Agency
Washington, DC 20460**

DISCLAIMER

This document has been reviewed in accordance with United States Environmental Protection Agency policy and approved for publication and distribution. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

TABLE OF CONTENTS

FOREWORD Page iv

ACKNOWLEDGMENTS Page vi

A Summary of General Assessment Factors for Evaluating the Quality of Scientific
and Technical Information Page 1

1. Introduction Page 1

 1.1 Overview Page 1

 1.2 Purpose Page 1

 1.3 Background Page 2

2. Assessment Factors Page 4

 2.1 General Assessment Factors Page 4

 2.2 Examples of Questions Raised by Consideration of the
 Assessment Factors Page 5

 2.3 Relationship Between the General Assessment Factors and the
 Elements of Quality in EPA’s Information Quality Guidelines Page 8

3. Summary Page 10

References Page 11

FOREWORD

This document was prepared under the auspices of the Science Policy Council (SPC) to describe the assessment factors and considerations generally used by the Agency to evaluate the quality and relevance of scientific and technical information. These general assessment factors are founded in the Agency guidelines, practices and procedures that make up the EPA information and quality systems, including existing program-specific quality assurance policies. As such, the general assessment factors do not constitute new quality-related considerations, nor does this document describe a new process for evaluating information. This document is intended to raise the awareness of the information-generating public about EPA's ongoing interest in ensuring and enhancing the quality of information available for Agency use. Further, it complements the *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency* (EPA Information Quality Guidelines). This summary of Agency practice is also an additional resource for Agency staff as they evaluate the quality and relevance of information, regardless of source.

Consistent with the Agency's approach to the development of the EPA Information Quality Guidelines, this document is the product of an open, collaborative process between EPA and the public. During the development of this document, EPA obtained public comments on a draft version of the document released in September 2002 and commissioned the National Academy of Sciences to host a workshop in January 2003 to discuss key aspects of this document from a scientific and technical perspective.

We want to acknowledge and thank the Assessment Factors workgroup for its steady and insightful work in assembling this document under stringent time constraints and scrutiny. We particularly appreciate the efforts of the co-chairs, Halûk Özkaynak (ORD) and Greg Schweer (OPPTS), who successfully led and shepherded the workgroup.

It is with great pleasure that we present the *Summary of General Assessment Factors for Evaluating the Quality of Scientific and Technical Information*.

Paul Gilman, Ph.D.
Science Advisor to the Agency
Chair, Science Policy Council

Elaine Stanley
Director, Office of Information Analysis
and Access
Office of Environmental Information

Science Policy Council

Paul Gilman, Science Advisor to
the Agency, Chair
Devereaux Barnes, OSWER
Robert Brenner, OAR
Jerry Clifford, OIA
William Farland, ORD
Geoffrey Grubbs, OW
Diana Love, OECA
Bharat Mathur, Region 5
Albert McGartland, OPEI

Stan Meiburg, Region 4
Joseph Merenda, OPPTS
William Muszynski, Region 2
Joanne Rodman, OCHP
Michael Ryan, OCFO
Michael Shapiro, OW
Elaine Stanley, OEI
Ramona Trovato, OEI
Vanessa Vu, SAB
Anna Wolgast, OGC

Science Policy Council Steering Committee

Randolph Perfetti, Chair
Thomas Baugh, Region 4
Michael Brody, OCFO
Reginald Cheatham, OEI
Patricia Cirone, Region 10
John Diamante, OIA
Michael Firestone, OCHP
A. Robert Flaak, SAB
Jerri-Anne Garl, Region 5
Roland Hemmett, Region 2
Lee Hofmann, OSWER
Daniel Malloy, OCFO

Kate Mahaffey, OPPTS
Carl Mazza, OAR
James Nelson, OGC
Jennifer Orme-Zavaleta, ORD
Rosemarie Russo, ORD
Rita Schoeny, OW
Margaret Stasikowski, OPPTS
Kevin Teichman, ORD
Mary Ellen Weber, OPPTS
William Wood, ORD
Tracey Woodruff, OPEI

Science Policy Council Staff

Edward Bender

Kerry Dearfield

Kathryn Gallagher

ACKNOWLEDGMENTS

Many people worked on and contributed to the assessment factors effort that ultimately resulted in this document. Many EPA employees from the Agency's Offices and Regions provided input and we would like to specifically acknowledge the efforts and contributions made by the following individuals:

Halûk Özkaynak, ORD, Co-Chair
Mary Belefski, OPPTS
Katherine Biggs, OECA
Connie Bosma, ORD
Ming Chang, OEI
Weihsueh Chiu, ORD
Evangeline Cummings, OEI
Kerry Dearfield, SPC Staff
Thomas Eagles, OAR
Kathryn Gallagher, SPC Staff
Staci Gatica-Hebert, ORD
Roland Hemmett, Region 2
Annie Jarabek, ORD
Stephen Kroner, OSWER

Greg Schweer, OPPTS, Co-Chair
Kevin Kubik, Region 2
Karen Martin, OAR
Amy Mills, ORD
Barbara Pace, OGC
Nicole Paquette, OEI
Devon Payne-Sturges, OPEI
Ward Penberthy, OPPTS
James Quackenboss, ORD
Mary Reiley, OW
Bruce Rodan, ORD
Terry Simpson, ORD
Elaine Stanley, OEI
Lanelle Wiggins, OPEI

U.S. Environmental Protection Agency

A Summary of General Assessment Factors for Evaluating the Quality of Scientific and Technical Information

1. Introduction

1.1 Overview

As part of the ongoing commitment of the United States Environmental Protection Agency (USEPA) to ensure the quality of the information it uses, the Agency is publishing this summary of general assessment factors in an effort to enhance the transparency about EPA's quality expectations for information that is voluntarily submitted to or gathered or generated by the Agency for various purposes. This Assessment Factors document is intended to inform information-generating scientists about quality issues that should appropriately be taken into consideration at the time information is generated. It is also an additional resource for Agency staff as they evaluate the quality and relevance of information, regardless of source. The general assessment factors are drawn from the Agency's existing information quality systems, practices and guidelines that describe the types of considerations EPA takes into account when evaluating the quality and relevance of scientific and technical information used in support of Agency actions. As such, the general assessment factors do not constitute new quality-related considerations, nor does this document describe a new process for evaluating information. This document is intended to raise the awareness of the information-generating public about EPA's ongoing interest in ensuring and enhancing the quality of information available for Agency use.

1.2 Purpose

The Agency believes that the summary of general assessment factors provided in this document will serve to increase the extent to which the information-generating public builds quality considerations into the generation and documentation of their information products. The Agency expects that the resulting improvements in the quality of such information will enable the Agency to more fully utilize and disseminate such information. Thus, this document is intended to complement the *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency* (EPA Information Quality Guidelines) (EPA, 2002) and other Agency efforts to ensure and enhance information quality, as discussed below in Section 1.3. This document is not a regulation and is not intended to create any legal rights or impose legally binding requirements or obligations on EPA or the information-generating public.

Although the assessment factors as presented are intended to most generally apply to individual pieces of information, they can also be used as part of a broader evaluation of a body of evidence that is collectively evaluated through a process typically referred to as a “weight-of-evidence” approach. The weight-of-evidence approach considers all relevant information in an integrative assessment that takes into account the kinds of evidence available, the quality and quantity of the evidence, the strengths and limitations associated with each type of evidence and explains how the various types of evidence fit together. Details as to the Agency’s approach to integrating a body of evidence depend on the type of decision or action being undertaken, and are not the subject of this document. For instance, the *Guidelines for Carcinogen Risk Assessment, Review Draft* (EPA, 1999) provides guidance on characterizing the weight-of-evidence for carcinogenicity. Similarly, the *Guidelines for Ecological Risk Assessment* (EPA, 1998) describes the development of “lines of evidence” to reach a conclusion regarding an ecological risk estimate.

The general assessment factors are presented and discussed more fully in Section 2.1. Section 2.2 presents illustrative examples of the types of questions that consideration of these factors raise in the process of evaluating the quality and relevance of different types of information for different uses. The relationship between these general assessment factors and the elements of quality contained in the EPA Information Quality Guidelines is discussed in Section 2.3.

1.3 Background

In October 2002, EPA made available the EPA Information Quality Guidelines. The EPA Information Quality Guidelines were developed in response to guidelines issued by the Office of Management and Budget (OMB, 2002) under Section 515 of the Treasury and General Government Appropriations Act for Fiscal Year 2001 (Public Law 106-554; H.R. 5658). The EPA Information Quality Guidelines set forth the Agency’s policy and procedural guidance for ensuring and maximizing the quality of information *disseminated* by EPA, regardless of the source of the information, and articulate the Agency’s ongoing commitment to ensuring and maximizing information quality through existing policies, systems and programs. Thus, the EPA Information Quality Guidelines build upon the Agency’s numerous existing systems, practices and guidelines that address information quality, and provide new policies and administrative mechanisms that respond to OMB’s guidelines.

The EPA Information Quality Guidelines also recognize that, as part of its efforts to ensure information quality, the Agency does not wait until the point at which information is *disseminated* to consider important quality principles. Rather, the Agency recognizes that it is

important to assure the quality of information through processes that incorporate quality principles starting at the point at which information is *generated*.

The Agency uses and disseminates information that is generated by a variety of sources, including EPA itself as well as other parties that produce information through EPA contracts, grants and cooperative and interagency agreements or in response to a requirement under a statute, regulation, permit, order or other mandate. EPA generally has considerable control or influence over the quality of this information at the time the information is generated. Existing quality controls that EPA applies to the generation of information from these sources are based on EPA's Quality System (EPA, 2000a; EPA, 2000b), Peer Review Policy (EPA, 1994), Risk Characterization Policy (EPA, 1995) and other agency-wide and program-specific policies, as well as specific provisions in contracts, grants, agreements, regulations and statutes. A few additional useful web sites for obtaining further information on EPA's Quality System and various regulatory policies and decisions are provided under the References section at the end of this document.

The Agency also receives information that is voluntarily submitted by or collected from external sources, the generation of which does not come under the direct control of the Agency's internal information quality systems. This information may include scientific studies published in journal articles, testing or survey data, such as environmental monitoring or laboratory test results, and analytic studies, such as those that model environmental conditions or that assess risks to public health. Since EPA has placed great emphasis on the management of environmental issues on a cooperative basis with its many stakeholders, the amount of information submitted to EPA by external sources is increasing. Such sources include other federal, state, tribal, local and international agencies; national laboratories; academic and research institutions; business and industry; and public interest organizations. Although EPA's existing quality systems are not applied at the time this information is generated, EPA does apply appropriate quality controls when evaluating this information for use in Agency actions and for its dissemination consistent with the EPA Information Quality Guidelines. The Agency hopes this document will inform the public of EPA's objectives and enlist them in its effort to disseminate quality information and make quality decisions.

During the development of this document, EPA requested public input in a variety of ways. EPA distributed a draft document for public comment in September 2002 and hosted a public meeting in Washington, DC. In January 2003, EPA commissioned the National Academy of Sciences to host a workshop to discuss key aspects of this document from a scientific and technical perspective. EPA revised this document based on the input received through these public outreach opportunities.

2. Assessment Factors

2.1 General Assessment Factors

When evaluating the quality and relevance of scientific and technical information, the considerations that the Agency typically takes into account can be characterized by five general assessment factors:

- ***Soundness*** - *The extent to which the scientific and technical procedures, measures, methods or models employed to generate the information are reasonable for, and consistent with, the intended application.*
- ***Applicability and Utility*** - *The extent to which the information is relevant for the Agency's intended use.*
- ***Clarity and Completeness*** - *The degree of clarity and completeness with which the data, assumptions, methods, quality assurance, sponsoring organizations and analyses employed to generate the information are documented.*
- ***Uncertainty and Variability*** - *The extent to which the variability and uncertainty (quantitative and qualitative) in the information or in the procedures, measures, methods or models are evaluated and characterized.*
- ***Evaluation and Review*** - *The extent of independent verification, validation and peer review of the information or of the procedures, measures, methods or models.*

These assessment factors reflect the most salient features of EPA's existing information quality policies and guidelines. Whether the information consists of scientific theories, computer codes for modeling environmental systems, environmental monitoring data, economic analyses, social survey or demographic data, chemical toxicity testing, environmental fate and transport predictions or a human health risk assessment, EPA generally evaluates information by weighing considerations that fit within these five assessment factors. Thus, these factors encompass considerations that are weighed in the process of evaluating the quality and relevance of information. The appropriate level of quality for any particular information product is necessarily related to how and in what context the information is to be used. If EPA chooses to later "disseminate" the information, that dissemination would be covered by the Information Quality Guidelines which describe EPA policy and procedures for reviewing and substantiating the quality of information before EPA disseminates it.

When EPA considers using information for a particular purpose, careful judgment is applied to evaluate the information for quality and relevance in the context of the specific Agency action being developed. For instance, in the context of a given action, EPA may need to weigh the appropriateness of using information with significant, but known uncertainties to fill “data gaps,” relative to using default assumptions or committing additional resources to generate new information.

2.2 Examples of Questions Raised by Consideration of the Assessment Factors

Example questions that could be raised by the consideration of each of the assessment factors for various types of information are provided below. Given the very general nature of these assessment factors, the Agency felt that a compilation of such illustrative questions would most clearly convey the intended nature and breadth of the assessment factors, and how they would be reflected in an evaluation of various types of information. However, the applicability of these factors depends on the individual situation, and EPA retains discretion to consider and use factors and approaches on a case-by-case basis that may differ from the illustrative considerations presented below.

2.2.1 Soundness

The extent to which the scientific and technical procedures, measures, methods or models employed to generate the information are reasonable for, and consistent with, the intended application.

- a) Is the purpose of the study reasonable and consistent with its design?
 - b) To what extent are the procedures, measures, methods, or models employed to develop the information reasonable and consistent with sound scientific theory or accepted approaches?
 - c) How do the study’s design and results compare with existing scientific or economic theory and practice? Are the assumptions, governing equations and mathematical descriptions employed scientifically and technically justified? Is the study based on sound scientific or econometric principles?
 - d) In the case of a survey, have the questionnaires and other survey instruments been validated (e.g., compared with direct measurement data)? Were checks for potential errors made during the interview process?
-

- e) How internally consistent are the study's conclusions with the data and results presented?

2.2.2 *Applicability and Utility*

The extent to which the information is relevant for the Agency's intended use.

- a) How useful or applicable is the scientific or economic theory applied in the study to the Agency's intended use of the analysis?
- b) How relevant are the study's purpose, design, outcome measures and results to the Agency's intended use of the analysis (e.g., for a chemical hazard characterization)?
- c) Are the domains (e.g., duration, species, exposure) where the model or results are valid useful to the Agency's application?
- d) How relevant is the study to current conditions of interest? For example, in the case of a survey, are conditions likely to have changed since the survey was completed (i.e., is the information still relevant)? Is the sampled population relevant to the Agency's current application? How well does the sample take into account sensitive subpopulations?

2.2.3 *Clarity and Completeness*

The degree of clarity and completeness with which the data, assumptions, methods, quality assurance, sponsoring organizations and analyses employed to generate the information are documented.

- a) To what extent does the documentation clearly and completely describe the underlying scientific or economic theory and the analytic methods used?
 - b) To what extent have key assumptions, parameter values, measures, domains and limitations been described and characterized?
 - c) To what extent are the results clearly and completely documented as a basis for comparing them to results from other similar tests?
-

- d) If novel or alternative theories or approaches are used, how clearly are they explained and the differences with accepted theories or approaches highlighted?
- e) Is the complete data set accessible, including metadata, data-dictionaries and embedded definitions (e.g., codes for missing values, data quality flags and questionnaire responses)? Are there confidentiality issues that may limit accessibility to the complete data set?
- f) In the case of a modeling exercise, have the definitions and units of model parameters been provided? To what extent have the procedures for applying the model been clearly and completely documented? How available and adequate is the information necessary to run the model computer code?
- g) To what extent are the descriptions of the study or survey design clear, complete and sufficient to enable the study or survey to be reproduced?
- h) Have the sponsoring organization(s) for the study/information product and the author(s) affiliation(s) been documented?
- i) To what extent are the procedures for quality assurance and quality control of the data documented and accessible?

2.2.4 *Uncertainty and Variability*

The extent to which the variability and uncertainty (quantitative and qualitative) in the information or in the procedures, measures, methods or models are evaluated and characterized.

- a) To what extent have appropriate statistical techniques been employed to evaluate variability and uncertainty? To what extent have the sensitive parameters of models been identified and characterized?
 - b) To what extent do the uncertainty and variability impact the conclusions that can be inferred from the data and the utility of the study? What are the potential sources and effects of error and bias in the study design?
-

- c) Did the study identify potential uncertainties such as those due to inherent variability in environmental and exposure-related parameters or possible measurement errors?

2.2.5 Evaluation and Review

The extent of independent verification, validation and peer review of the information or of the procedures, measures, methods or models.

- a) To what extent has there been independent verification or validation of the study method and results? What were the conclusions of these independent efforts, and are they consistent?
- b) To what extent has independent peer review been conducted of the study method and results, and how were the conclusions of this review taken into account?
- c) Has the procedure, method or model been used in similar, peer reviewed studies? Are the results consistent with other relevant studies?
- d) In the case of model-based information, to what extent has independent evaluation and testing of the model code been performed and documented?

2.3 Relationship Between the General Assessment Factors and the Elements of Quality in EPA's Information Quality Guidelines

The definition of quality in the EPA Information Quality Guidelines consists of three components, consistent with the definition of quality in OMB's Guidelines: *objectivity*, *utility* and *integrity* of disseminated information. "Objectivity" focuses on the extent to which information is presented in an accurate, clear, complete and unbiased manner; and, as a matter of substance, the extent to which the information is accurate, reliable and unbiased. "Utility" refers to the usefulness of the information to the intended users. "Integrity" refers to security, such as the protection of information from unauthorized access or revision, to ensure the information is not compromised through corruption or falsification.

The five general assessment factors presented in this document are consistent with the quality elements of *objectivity* and *utility*, but do not extend to the distinct element of *integrity* (which refers to the separate matter of security issues). The assessment factor *applicability and utility* is most directly related to the element of *utility* in the OMB and EPA Information Quality

Guidelines. The other four assessment factors relate to the element of *objectivity*, which itself encompasses a number of issues related to both presentation and substance. In particular, the factor *clarity and completeness* is most directly related to some aspects of the presentation of information (including whether the information is “presented in an accurate, clear, complete and unbiased manner”). The factors *soundness, uncertainty and variability* and *evaluation and review* most directly relate to the substantive aspects of the element of *objectivity* (related to whether the information itself is “accurate, reliable and unbiased”), although they also play a role in enhancing aspects of the presentation of the information. Thus, the general assessment factors are fully consistent with the related information quality elements described in the OMB and EPA Information Quality Guidelines, and do not constitute a conceptually different or unrelated basis for evaluating information quality.

It is important to note that the EPA Information Quality Guidelines apply to “information” that EPA disseminates to the public. The EPA Information Quality Guidelines apply to information generated by third parties if EPA distributes information prepared or submitted by an outside party in a manner that reasonably suggests that EPA endorses or agrees with it; if EPA indicates in its distribution that the information supports or represents EPA's viewpoint; or if EPA in its distribution proposes to use or uses the information to formulate or support a regulation, guidance, policy or other Agency decision or position (EPA 2002). Please refer to the EPA Information Quality Guidelines for additional information regarding their applicability to information EPA disseminates.

3. Summary

This document describes the assessment factors and considerations generally used by the Agency to evaluate the quality and relevance of the broad range of scientific and technical information used by the EPA. These factors are founded in the Agency guidelines, practices and procedures that make up the EPA information and quality systems including existing program-specific quality assurance policies. Consistent with the Agency's approach to the development of the EPA Information Quality Guidelines, this document is the product of an open, collaborative process between EPA and the public.

The Agency believes that the summary of general assessment factors provided in this document will serve to increase the extent to which the information-generating public builds quality considerations into the generation and documentation of their information products. The Agency expects that the resulting improvements in the quality of such information will enable the Agency to more fully utilize and disseminate such information. Thus, this document is intended to complement the EPA Information Quality Guidelines and other Agency efforts to ensure and enhance information quality.

References

- US Environmental Protection Agency, 1994, *Peer Review and Peer Involvement at the U.S. Environmental Protection Agency*, Science Policy Council, <http://www.epa.gov/osp/spc/memo0607.htm>
- US Environmental Protection Agency, 1995, *Policy for Risk Characterization*, Science Policy Council, <http://www.epa.gov/osp/spc/rcpolicy.htm>
- US Environmental Protection Agency, 1998, *Guidelines for Ecological Risk Assessment*, EPA/630/R095/002F, Risk Assessment Forum, <http://cfpub.epa.gov/ncea/cfm/recordisplay.cfm?deid=12460>
- US Environmental Protection Agency, 1999, *Guidelines for Carcinogen Risk Assessment, Review Draft*, NCEA-F-0644, Office of Research and Development, <http://cfpub.epa.gov/ncea/raf/cancer.cfm>
- US Environmental Protection Agency, 2000a, "Policy and Program Requirements for the Mandatory Agency-wide Quality System," (EPA Order 5360.1 A2), <http://www.epa.gov/quality/qs-docs/5360-1.pdf>
- US Environmental Protection Agency, 2000b, "EPA Quality Manual for Environmental Programs," (EPA Order 5360 A1), <http://www.epa.gov/quality/qs-docs/5360.pdf>
- US Environmental Protection Agency, 2002, *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency*, EPA 260R-02-008, Office of Environmental Information, <http://www.epa.gov/oei/qualityguidelines>
- US Office of Management and Budget, 2002, *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies*, (67 FR 8452), <http://www.whitehouse.gov/omb/fedreg/reproducible2.pdf>

Additional Useful Web Sites

- EPA Quality System web site: <http://www.epa.gov/quality>
- EPA Science Policy Council web site: <http://www.epa.gov/osp/spc>
- EPA Information Quality Guidelines web site: <http://www.epa.gov/oei/qualityguidelines>
-



December 2012

U.S. Environmental Protection Agency

Guidance for Evaluating and Documenting the Quality of Existing Scientific and Technical Information

*Addendum to: A Summary of General Assessment Factors for
Evaluating the Quality of Scientific and Technical Information*

**Prepared for the U.S. Environmental Protection Agency
by members of the Peer Review Advisory Group
a group of the EPA's Science and Technology Policy Council**

**Science and Technology Policy Council
U.S. Environmental Protection Agency
Washington, DC 20460**

This page is intentionally left blank

**Guidance for Evaluating and Documenting the Quality of
Existing Scientific and Technical Information**
**Addendum to: *A Summary of General Assessment Factors for Evaluating the Quality
of Scientific and Technical Information***

1. Overview

In 2010, the U.S. Environmental Protection Agency's (EPA or Agency) Office of Inspector General (OIG) reviewed the process used by EPA to support its greenhouse gases endangerment finding (EPA, 2009a). The OIG's findings were published in the report, *Procedural Review of EPA's Greenhouse Gases Endangerment Finding Data Quality Processes* (EPA, 2011a). The report recommended that the Agency revise its guidance document, *A Summary of General Assessment Factors for Evaluating the Quality of Scientific and Technical Information* (EPA, 2003), "to establish minimum review and documentation requirements for assessing and accepting data from other organizations." This Addendum responds to the OIG's recommendation by providing guidance for assessing and accepting existing scientific and technical information. It is relevant not only to data from other organizations, but to any existing scientific and technical information used to support Agency decision making.

This Addendum contains guidance for:

- assessing and accepting existing scientific and technical information, and
- documenting the review and analysis of existing scientific and technical information.

The Addendum also contains illustrative examples of approaches for applying the guidance.

2. Background

EPA uses and disseminates scientific and technical information obtained from a variety of sources, both internal and external. Information generated by the Agency, or obtained through EPA contracts, grants, and cooperative and interagency agreements, falls under the direct control of the Agency's internal information quality systems and various Agency-wide and program-specific policies and procedures (EPA, 1994; EPA, 2002; EPA, 2006; EPA, 2008a,b; EPA, 2009b; EPA, 2011b; EPA, 2012a,b). Information generated by or obtained from outside sources, such as local and state governments, tribes, industry, environmental organizations, other federal agencies, and the peer-reviewed literature, is evaluated by EPA using the guidance contained in the following documents to determine whether it meets the quality requirements of the Agency:

- *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency* (EPA, 2002);

- *A Summary of General Assessment Factors for Evaluating the Quality of Scientific and Technical Information* (the document to which this Addendum applies; EPA, 2003);
- *Guidance on Quality Assurance Project Plans* (EPA, 2012c); and
- *Peer Review Handbook, 3rd Edition* (EPA, 2006) and its Addendum (EPA, 2009c).

Sometimes, information may be used for purposes other than those for which they were originally intended. An example is the use of historical municipal drinking water data in Agency studies of groundwater contamination. Another example is the use of information derived from the scientific literature, such as epidemiologic or experimental studies from peer reviewed journals identified in the PubMed or ToxNet databases. Such information is referred to as “existing” data or information. Regardless of its source, this information should be evaluated to verify that its quality is appropriate for its intended use by the Agency (EPA, 2002).

3. Guidance for Evaluating and Documenting Existing Scientific and Technical Information

3.1. Assessing and Accepting Existing Scientific and Technical Information

When collecting and assessing existing scientific and technical information, use the five general assessment factors (Soundness, Applicability and Utility, Clarity and Completeness, Uncertainty and Variability, and Evaluation and Review) found in the Assessment Factors guidance document (EPA, 2003) to determine whether the information complies with EPA’s Information Quality Guidelines (EPA, 2002). Sample questions for evaluating the quality of the information are offered in Section 2.2 of the Assessment Factors guidance document. Refer to the peer review considerations found in Section 2.2.17 of the Agency’s Peer Review Handbook, 3rd Edition (EPA, 2006) for help in addressing the “Evaluation and Review” factor. Section 2.2.17 of the Handbook states that scientific and technical work products important to EPA environmental decision making are candidates for peer review, regardless of whether they were produced by the Agency or by an outside organization. Often, the existing information has already undergone independent peer review, and in such cases, the review should meet the intent of the Agency’s peer review policy and be commensurate with EPA’s proposed use of the information.

The criteria for accepting existing information (called acceptance or performance criteria) should be tailored to the type of information under consideration based on the principle of a “graded approach,” in which the level of quality assurance applied to the information is commensurate with the intended use of the information and the degree of confidence necessary in that information (EPA, 2002). A full discussion of acceptance criteria may be found in the guidance *Handbook for Developing Quality Assurance Project Plans* (EPA, 2012c), which includes definitions in Appendix B for six data quality indicators (Precision, Bias,

Representativeness, Completeness, Comparability, and Sensitivity, or PBRCCS) considered important to environmental studies.

Examples of the use of the five assessment factors, a graded approach, and the application of some acceptance criteria may be found in Section 4 of this Addendum.

3.2. Documenting the Review and Analysis of Existing Scientific and Technical Information

EPA organizations are expected to develop and use a Quality Assurance Project Plan (QAPP), or an equivalent form of documentation, to document the procedures used in the review and analysis of existing scientific and technical information. Such documentation is part of EPA's mandatory Quality Program (EPA, 2012c; see Chapter 3 for relevance to existing data). The QAPP or its equivalent should include a description of the type and quality of information needed for a specific decision or use, it should establish the acceptance criteria or quality determinations against which the information will be evaluated, and it should document the review and analysis process for the 5 assessment factors. And finally, the QAPP should describe how the outcomes (or results) of the review and analysis process will be documented and reported. The graded approach applies as well to documentation; i.e., the level of effort expended to document the review and analysis process should be commensurate with the intended use of the information and the degree of confidence required. The requirements for a QAPP may be found in Annex B of CIO 2106-S-01 (for EPA organizations) and CIO 2106-S-02 (for non-EPA organizations) (EPA, 2012a,b).

A checklist of QAPP elements that may be applied during documentation is provided in Annex B of *Handbook for Developing Quality Assurance Project Plans* (EPA, 2012c), and examples of documentation, both simple and detailed, are included in Section 4 below.

4. Examples

The following examples have been included for illustrative purposes only to demonstrate how the five assessment factors can be reviewed and documented. They indicate a range of options and different levels of complexity, taking into account the graded approach. Users may adapt these examples as models for developing their own quality review and documentation of the assessment factors. Note that this process may be one piece of an overall evaluation for deciding whether to accept or reject existing data or information.

Quality Assurance Project Plan (QAPP) Template

The following QAPP template was developed by ORD's National Center for Environmental Assessment (NCEA). It has been shared with scientists in the Agency and with EPA contractors as a model for developing a QAPP for conducting a literature search and analyzing the quality of existing studies. General guidance and a checklist for evaluating key studies are included.

Elements of a Quality Assurance Project Plan (QAPP) For Collecting, Identifying, and Evaluating Existing Data/Information

<http://www2.epa.gov/osa/elements-quality-assurance-project-plan-qapp-collecting-identifying-and-evaluating-existing>

Sample Quality Assurance Project Plans (QAPPs)

The following QAPP example is intended to be applicable to both existing data as well as existing literature. An example for documenting the evaluation of the five assessment factors is included as Appendix 1: Reference Evaluation Template.

Data and Literature Evaluation for the EPA's Study of the Potential Impacts of Hydraulic Fracturing (HF) on Drinking Water Resources

<http://www2.epa.gov/hfstudy/qapp-revision-no-2-data-and-literature-evaluation-epas-study-potential-impacts-hydraulic>

The following QAPP illustrates the use of the graded approach in planning and documenting a data collection study based on the compilation and use of existing data. In order to assess and report on the ecological health of the NJ-NY Harbor Estuary, the New England Interstate Water Pollution Control Commission (NEIWPCC) developed this QAPP to describe the activities needed to identify and evaluate existing data used in the final report. Historically, the term "secondary data" used in this QAPP was interchangeable with the term "existing data".

State of the Estuary Report QAPP

http://www.epa.gov/region1/measure/qapp_examples/pdfs/SOE-QAPP.pdf

The following example illustrates how the quality and relevance of existing information can be evaluated for use by reviewing and documenting the assessment factors.

Illustrative Example for Applying Assessment Factors in Collecting, Identifying and Evaluating Existing Literature

<http://www2.epa.gov/osa/equivalent-quality-assurance-project-plan-qapp-illustrative-example-review-and-documentation>

Sample Checklist

The following checklist, used by Region 10, considers three criteria that existing information should demonstrate: traceability, accessibility, and documentation. Traceability provides the original source and publication information; accessibility gives the exact location and format of the information; and documentation provides information to support legal scrutiny covering *quality, usability, integrity, objectivity and reproducibility*.

Checklist for the Assessment of Existing Information

<http://www2.epa.gov/osa/checklist-assessment-existing-informationsecondary-data>

Criteria and Evaluations – Overall Process Example

The following example provided by the IRIS Program illustrates how the assessment factors can be applied to searches for mechanistic evidence in published data and information as part of a broad evaluation of a body of evidence—also referred to as a weight-of-evidence approach.

Defining Assessment Factors (e.g., exclusion/inclusion criteria)

<http://www2.epa.gov/osa/defining-assessment-factors>

The following table, also provided by the IRIS program, demonstrates the documentation of an analysis of findings in the scientific literature. It incorporates the process for evaluating and accepting information in the context of the overall project.

Evaluating the Quality of Individual Studies

<http://www2.epa.gov/osa/evaluating-quality-individual-studies>

References

- Office of Management and Budget, *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies*, Federal Register / Vol. 67, No. 36 / Tuesday, February 22, 2002, p. 8452.
- US Environmental Protection Agency (EPA), 1994, *Peer Review and Peer Involvement at the U.S. Environmental Protection Agency*, Science Policy Council, <http://www2.epa.gov/osa/products-and-publications-related-science-and-technology-produced-office-science-advisor>
- US Environmental Protection Agency, 2002, *Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by the Environmental Protection Agency*, EPA/260R-02-008, Office of Environmental Information, http://www.epa.gov/quality/informationguidelines/documents/EPA_InfoQualityGuidelines.pdf
- US Environmental Protection Agency, 2003, *A Summary of General Assessment Factors for Evaluating the Quality of Scientific and Technical Information*, EPA 100/B-03/001, Science Policy Council, <http://www2.epa.gov/osa/summary-general-assessment-factors-evaluating-quality-scientific-and-technical-information>
- US Environmental Protection Agency, 2006, *Peer Review Handbook, 3rd Edition*, EPA/100/B-06/002, Science Policy Council, <http://www2.epa.gov/osa/peer-review-handbook-3rd-edition-2006-and-addendum>
- US Environmental Protection Agency, 2008a, "EPA Quality Policy," (EPA Order CIO 2106.0), <http://www.epa.gov/irmpoli8/policies/21060.pdf>
- US Environmental Protection Agency, 2008b, "EPA Procedure for Quality Policy," (EPA Order CIO 2106-P-01.0), <http://www.epa.gov/irmpoli8/policies/2106p01.pdf>
- US Environmental Protection Agency, 2009a, Endangerment and Cause or Contribute Findings for Greenhouse Gases Under Section 202(a) of the Clean Air Act; Final Rule, Federal Register / Vol. 74, No. 239 / December 15, 2009, p. 66496.
- US Environmental Protection Agency, 2009b, *Guidance for Labeling Externally Validated Laboratory Analytical Data for Superfund Use*, EPA-540-R-08-005, Office of Solid Waste and Emergency Response, <http://www.epa.gov/superfund/policy/pdfs/EPA-540-R-08-005.pdf>
- US Environmental Protection Agency, 2009c, *Addendum to the Peer Review Handbook, 3rd Edition: Appearance of a Lack of Impartiality in External Peer Reviews*, Science Policy Council, <http://www2.epa.gov/osa/peer-review-handbook-3rd-edition-2006-and-addendum>

US Environmental Protection Agency, 2011a, *Procedural Review of EPA's Greenhouse Gases Endangerment Finding Data Quality Processes*, Report No. 11-P-0702, Office of Inspector General, <http://www.epa.gov/oig/reports/2011/20110926-11-P-0702.pdf>

US Environmental Protection Agency, 2011b, *EPA's Action Development Process: Guidance for EPA Staff on Developing Quality Actions*, Office of Policy, [http://yosemite.epa.gov/sab/sabproduct.nsf/5088B3878A90053E8525788E005EC8D8/\\$File/adp03-00-11.pdf](http://yosemite.epa.gov/sab/sabproduct.nsf/5088B3878A90053E8525788E005EC8D8/$File/adp03-00-11.pdf)

US Environmental Protection Agency, 2012a, "DRAFT FINAL EPA Quality Standard for Environmental Data Collection, Production and Use by EPA Organizations," (EPA Order CIO 2106-S-01.0)

US Environmental Protection Agency, 2012b, "DRAFT FINAL EPA Quality Standard for Environmental Data Collection, Production, and Use by Non-EPA (External) Organizations," (EPA Order CIO 2106-S-02.0), <http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OEI-2012-0774-0021>

US Environmental Protection Agency, 2012c, "DRAFT FINAL Handbook for Developing Quality Assurance Project Plans," <http://www.regulations.gov/#!documentDetail;D=EPA-HQ-OEI-2012-0774-0003>, (as of Nov. 30, 2012, the link will be <http://intranet.epa.gov/quality/drfaqdocuments.html>)

A Systematic Approach for Evaluating the Quality of Experimental Toxicological and Ecotoxicological Data¹

H.-J. KLIMISCH,² M. ANDREAE, AND U. TILLMANN

BASF Aktiengesellschaft, D-67056 Ludwigshafen, Germany

Received November 7, 1996

The evaluation of the quality of data and their use in hazard and risk assessment as a systematic approach is described. Definitions are proposed for reliability, relevance, and adequacy of data. Reliability is differentiated into four categories. Criteria relating to international testing standards for categorizing reliability are developed. A systematic documentation of evaluating reliability especially for use in the IUCLID database is proposed. This approach is intended to harmonize data evaluation processes worldwide. It may help the expert in subsequent assessments and should increase the clarity of evaluation. © 1997 Academic Press

INTRODUCTION

Hazard and risk assessment for "existing substances" must be carried out in Europe based on Council Regulation 793/93 (EEC, 1993) and following principles of Commission Regulation 1488/94 (EEC, 1994). All relevant available information/data and corresponding study reports of substances, published in a priority list, must be submitted by the manufacturer/importer using a special software package on disk (IUCLID: International Uniform Chemical Information Database) and as hard copies. During the risk assessment process the assessor must consider whether the supplied data are complete and valid for use in risk assessment. This is particularly important for data on "existing substances" (EINECS, 1981). There may be a number of test results available for each end point but some or all of them may have not been carried out following current standards in toxicology and ecotoxicology.

Before a hazard identification may be performed, the

supplied data must be evaluated considering their quality and adequacy for a risk assessment. Some general guidelines on data evaluation were published by the European Commission in a "Technical Guidance Document" (EU, 1994, 1995), based on general principles for data evaluation of the International Coordination of Criteria Document Production (IPCS, 1993). Considerations of the assessment on the quality of data have been described also by OECD (1994). Some experience in the evaluation of data was developed in Germany by the "GDCH-Advisory Committee on Existing Chemicals of Environmental Relevance (BUA)." During the past decade approximately 200 BUA reports were published by this committee composed of representatives of academia, the chemical industry, and government. During this work a systematic approach on data validation was found useful. It gives definitions, discusses a score system with different reliability categories according to validity, defines criteria, and generates a system for standardized documentation of validity evaluation to be used also in data sheets (IUCLID). It appeared to be useful to describe this approach on behalf of BUA in order to initiate harmonization of similar processes in data evaluation worldwide and to facilitate the exchange of experience toward improvement of such approaches. A characterization of the validity of experimental data should also help the expert to assess the effect of end points consistently and thus to increase clarity in hazard or risk assessment processes.

DEFINITIONS

Different terms are being used synonymously to characterize the quality of the data of toxicological and ecotoxicological studies: validation/validity, reliability, adequacy. These terms describe not only procedures to define the quality of test results (data), but also test methods are validated to prove their relevance and reproducibility. Validity statements are also given in assessment processes especially if an expert must decide which data of, for instance, conflicting study results are representative/relevant to describe an effect correctly.

¹ This paper is published on behalf of BUA (GDCH Advisory Committee on Existing Chemicals of Environmental Relevance, Federal Republic of Germany).

² To whom reprint requests should be addressed at Department of Toxicology.

The following definitions are proposed here to be used in hazard and risk assessment processes:

Reliability—Evaluating the inherent quality of a test report or publication relating to preferably standardized methodology and the way that the experimental procedure and results are described to give evidence of the clarity and plausibility of the findings.

Relevance—Covering the extent to which data and/or tests are appropriate for a particular hazard identification or risk characterization.

Adequacy—Defining the usefulness of data for risk assessment purposes. When there is more than one set of data for each effect, the greatest weight is attached to the most reliable and relevant.

The evaluation needs expert judgment and should be clear, so that the use made of a particular data set is clearly justified and understood by others. Agreement on standardized criteria for characterizing and differentiating the quality of data (their reliability, relevance, and adequacy) may be useful for a broader understanding and acceptance worldwide. Such evaluation of the quality of individual studies/data is a step in compiling data in the form of a "data sheets" for a substance (IUCLID, etc.) for hazard or risk assessment purposes. Such data sheets have the intention of making available all toxicological and ecotoxicological data about a substance and keeping them updated to the actual state of knowledge. Furthermore, if information about the quality of the individual test/data is given in such a data sheet, this would help to identify more easily those data preferably used for risk assessment.

CATEGORIES OF RELIABILITY

Test data of toxicological and ecotoxicological laboratory studies may be available as described in

- individual test reports
- publications (literature)
- review articles
- abstracts of presentations
- any other short information (safety data sheets, handbooks, etc.).

The more that details on methodology, test procedures, and analytics are documented, the easier an evaluation of their reliability should be in general. The amount of information presented will thus provide the basis for deciding on the reliability of data reported. Tests conducted and reported according to internationally accepted test guidelines (EU, EPA, FDA, OECD) and in compliance with the principles of Good Laboratory Practice (GLP) should have the highest grade of reliability and should be used as reference standards when evaluating the reliability of tests generated prior to the requirements of GLP and the international standardization of testing methods.

Our approach proposes to indicate a measure of the study/data reliability. Therefore, the quality of laboratory studies and of data from the literature may be differentiated and thus classified according to four categories of reliability.

The following categories/codes of reliability seem to be adequate:

Code	Category
1	Reliable without restriction
2	Reliable with restrictions
3	Not reliable
4	Not assignable

An additional Code 5 may be added to identify information/data which were *not evaluated* according to their reliability (special studies on, for instance, pharmacologic or mechanistic effects) without particular relevance for hazard/risk assessment.

The following definitions of these categories were found practicable to differentiate reliability (Codes 1–4):

1. *Reliable without Restriction*

This includes studies or data from the literature or reports which were carried out or generated according to generally valid and/or internationally accepted testing guidelines (preferably performed according to GLP) or in which the test parameters documented are based on a specific (national) testing guideline (preferably performed according to GLP) or in which all parameters described are closely related/comparable to a guideline method.

2. *Reliable with Restrictions*

This includes studies or data from the literature, reports (mostly not performed according to GLP), in which the test parameters documented do not totally comply with the specific testing guideline, but are sufficient to accept the data or in which investigations are described which cannot be subsumed under a testing guideline, but which are nevertheless well documented and scientifically acceptable.

3. *Not Reliable*

This includes studies or data from the literature/reports in which there are interferences between the measuring system and the test substance or in which organisms/test systems were used which are not relevant in relation to the exposure (e.g., unphysiologic pathways of application) or which were carried out or generated according to a method which is not acceptable, the documentation of which is not sufficient for an

assessment and which is not convincing for an expert judgment.

4. Not Assignable

This includes studies or data from the literature, which do not give sufficient experimental details and which are only listed in short abstracts or secondary literature (books, reviews, etc.).

CRITERIA FOR RELIABILITY CATEGORIES

In order to help in assigning a study to a category/code of reliability, some criteria should be considered more specifically, according to which the quality of the study in relation to standard methods and the scope of the documentation are assessed. Depending on the type of study, a differentiated evaluation by the expert is required: In the case of acute studies, the requirements may generally be interpreted more flexible and broadly than, for example, in the case of carcinogenicity studies. The following general criteria should be considered.

The standard methods recommended, e.g., by OECD or EU, are used as a reference. GLP principles should preferably be considered so that the reproducibility and acceptance according to the state-of-the-art of the results are guaranteed as far as possible. If a complete report is available or if the test, although not performed according to national/international standard methods, is described sufficiently and carried out according to a scientifically acceptable standard, the studies may be assessed as "reliable" as well. This also applies to literature publications. The basic data (test organisms, data on the method, and on the scope of the investigations) should be available and documented in the data set of the substance especially if a standard method was not used. Data on the purity of a substance are necessary particularly if impurities may have a substantial influence on the toxicity. This can be assessed only on a case-to-case basis. Information on dose/concentration is essential. Even if some criteria of an international standard are not met, the expert may decide that the study is "reliable with restrictions" and may be used for a risk assessment.

Toxicity Studies

The following information/data should generally be available and reported for *animal studies* which were not carried out according to an international/national standard method:

- Data/information on the test animals (species, strain, sex, age);
- Purity/composition/origin of the test substance;
- Number of animals evaluated;
- Scope of the investigations per animal (for instance, clinical chemistry, hematology, organ weights,

pathology or histopathology) and description of the methods;

- Description of the changes/lesions observed;
- Control group or historical control data of the laboratory;
- Description of the test conditions;
- Description of the route and doses of administration (preferably including analytical verification);
- Dose/concentration relationship if possible.

The following data/information should be available for *in vitro studies* which were not carried out according to an international/national standard method:

- Description of the test system and test method in details;
- Purity/composition/origin of the test substance;
- Data on the dose/concentration differentiated according to the toxicity of the test substance on the test system; information on volatility;
- Data on secondary effects which may influence a result (solubility, impurities, pH shifts, influence on the osmolarity, etc.);
- Appropriate negative/positive controls as integral parts of the test;
- References on adequacy of the method should be given or generally known.

The usefulness will be particularly influenced by the adequacy of the method.

Ecotoxicity Studies

For assessing the *reliability* of ecotoxicological studies, which are not carried out according to national/international test guidelines, the following items should be screened (expert judgment):

Acute studies.

- Clear description of the test procedure (complete documentation)
- Specification of the test substance (purity, by-products)
 - Data on the test species and the number of individuals tested
 - Data on the measured parameters (including definitions)
 - Data on exposure period
 - Use of emulgators/solubilizers³
 - Data on concentration control analysis³
 - Data on neutralization of samples⁴
 - Data on physical and chemical test conditions (pH value, conductivity, light intensity, temperature, hardness of water)

³ Especially in case of poorly soluble and unstable substances.

⁴ In case of basic and acid substances.

- Determined effect concentrations (EC/LC/NOEC/LOEC)
- Data on the statistical evaluations (including method)
- Data on dosing the test substance (static, semistatic, flow through system).

Additional items in case of chronic studies.

- Information about the investigated period of the life cycle of the test animals
- Data on feeding of test animals.

DOCUMENTATION OF RELIABILITY CATEGORIES IN DATA SHEETS (IUCLID)

A short justification should be given in writing for assigning data of a study to a code/category of reliability. This should help in making such an expert decision transparent and understandable. For codes/categories 1 and 2 only short phrases may be necessary to justify such an assignment: for instance, "OECD Guideline study: GLP," etc. A more detailed justification should be given particularly for studies which are assigned to Code 3 (unreliable). The justification must be documented. If the data are compiled in a data sheet (IUCLID), this may preferably be reported in such a computer-based system in an additional field "reliability" under each individual test. The responsible "European Chemicals Bureau" has generated a software package for the latest IUCLID version 2.12 (ECB, 1996) according to the following patterns:

Example: Reliability, Code number (wording) justification statement.

Code/Category of Reliability

Reliability 1. (Reliable without restriction) short free text phrases, for instance:

- Guideline study (OECD, etc.)
- Comparable to guideline study
- Test procedure according to national standards (DIN, etc.).

Reliability 2. (Reliable with restrictions) short free text, for instance:

- Acceptable, well-documented publication/study report which meets basic scientific principles
- Basic data given: comparable to guidelines/standards
- Comparable to guideline study with acceptable restrictions.

Reliability 3. (Not reliable) more detailed free text.

- Method not validated
- Documentation insufficient for assessment

— Does not meet important criteria of today standard methods

- Relevant methodological deficiencies
- Unsuitable test system.

Reliability 4. (Not assignable) short free text

- Only short abstract available
- Only secondary literature (review, tables, books, etc.).

Relevance/Adequacy

As described the evaluation of reliability is performed considering certain formal criteria using international standards as references. It should clearly be stated that it is not the intention of this procedure to automatically exclude all unreliable data from further consideration by experts in risk assessment. The classification into different reliability categories should help the assessor especially in cases when conflicting results regarding one end point are reported. In such cases results of studies with a higher reliability should have greater weight for being used in risk assessment.

If for example results of *in vitro* tests are available (positive and negative Ames test), the test with the higher reliability may be more relevant. Therefore the assessment of relevance is very important and only the expert can decide which test describes the effect "correctly." Ames tests carried out according to International Testing Guidelines and GLP but using different purities of the substance may lead to positive and negative results depending on the reactivity and quantity of impurities. Both tests may have a high reliability but only the test without the reactive impurity may be relevant if this is the chemical to be used. Only this test should be considered as adequate for risk assessment.

The *relevance* of an ecotoxicological study should be elucidated in the light of the following questions:

- Is the testing strategy (organism, exposure scenario) aligned with the occurrence and the persistence of the test substance in the environment (target compartment)?
- Is it possible to derive useful ecotoxicological information from data obtained from experiments with non-standard organisms (specialist, spread)?
- Are physical/chemical properties of the test substance (stability against hydrolytic and photolytic attacks, volatility, solubility) sufficiently considered before planning the test design?

Data with lower reliability may also be used as *supporting* information especially if the results are comparable or in the similar range; even in a case where only data with limited reliability are available, they may be used for definitive assessments of risk if the assessor considers these data as relevant (plausible) for risk assessment. For instance, LD₅₀ values from studies with

rats, rabbits, and dogs, each with limited information on methodology, were considered as of limited reliability or even unreliable. But despite these reliability limitations, the assessor may use such data for risk assessment if the LD₅₀ are within an acceptable range, evaluated in combination such that they are relevant (plausible), and show an only low interspecies variability. The same applies to a carcinogenicity study with too small a number of mice but showing a carcinogenic effect similar to that obtained in a reliable guideline study on rats.

It is not the aim of this paper to define criteria when studies with a restricted quality may nevertheless be used for hazard or risk assessment. This can only be decided by expert judgment on a case-by-case basis. Also data on structurally related compounds (SAR) should be used to define the relevance and adequacy of test results. All available experimental data as compiled in a data sheet (IUCLID) should be considered in risk assessment because only the totality of data will increase clarity of the conclusions. Thus limitations of publishing only a "definitive data set" appear to limit clarity and worldwide understanding. The relevance and adequacy of all the data used in a risk assessment process should be defined by expert judgment in a comprehensive report. Thus conclusions on relevance to humans of effects observed in studies in animals must be explained to make this interpretation clear and to

gain a better understanding of the mechanism of action of a substance.

The proposed systematic approach to define and differentiate reliability of data should help experts worldwide to decide about relevance of the data for humans and their adequacy in risk assessment processes.

REFERENCES

- EEC (1994). Commission Regulation No. 1488/94 of 28 June 1994 laying down the principles for the assessment of risks to man and the environment of existing substances. *Off. J. Eur. Communities* **37**(L 161), 3-11.
- EEC (1993). Council Regulation No. 793/93 of 22 March 1993 on the evaluation and control of the risks of existing substances. *Off. J. Eur. Communities* **36**(L 84), 1-7.
- EINECS (1981). European Inventory of Existing Commercial Chemical Substances (commercially available in the EEC between 01.01.71 until 18.09.81).
- EU (1994). European Commission, Directorate—General Environment, Nuclear Safety and Civil Protection: Risk Assessment of Existing Substances, Technical Guidance Document (XI, 919/94-EN).
- EU (1995). European Commission: Risk Assessment of New and Existing Substances; Technical Guidance Document Draft October.
- ECB (1996). Address of Dr. W. Karcher, CEC Joint Research Centre, European Chemicals Bureau, T.P. 280, I-21020 ISPRA (Varese) Italy; IUCLID: Additional Features of Version 2.12 (25.01.96).
- IPCS (1993). Meeting report on "International Co-ordination of criteria Document Production," Annex 5.
- OECD (1994). Revised Draft SIDS Manual (OECD Secretariat) EXCH, Manual 9405 DOC July.