epidemiologic studies, e.g., case control and cohort studies, to judge that some neurotoxic effect is associated with exposure. A case series in conjunction with other supporting evidence may also be judged "sufficient evidence." Epidemiologic and clinical case studies should discuss whether the observed effects can be considered biologically plausible in relation to chemical exposure. (Historically, much often has been made of the notion of causality in epidemiologic studies. Causality is a more stringent criterion than association and has become a topic of scientific and philosophical debate. See Susser [1986], for example, for a discussion of inference in epidemiology.)

Sufficient Experimental Animal Evidence/ Limited or No Human Data

This category includes agents for which there is sufficient evidence from experimental animal studies and/or limited human data to judge whether a potential neurotoxic hazard may exist. Generally, agents that have been tested according to current test guidelines would be included in this category. The minimum evidence necessary to judge that a potential hazard exists would be data demonstrating an effect in a single appropriate, well-executed study in a single experimental animal species. whereas the minimum evidence needed to judge that a potential hazard does not exist would include data from appropriate, wellexecuted laboratory animal studies that evaluated a variety of the potential manifestations of neuroxtoxicity and showed no effects at doses that were at least minimally toxic. Information on pharmacokinetics, mechanisms, or known properties of the chemical class may also strengthen the evidence.

Insufficient Evidence

This category includes agents for which there is less than the minimum evidence sufficient for identifying whether or not a neurotoxic hazard exists, such as agents for which there are no data on neurotoxicity or agents with data bases from studies in animals or humans that are limited by study design or conduct (e.g., inadequate conduct or report of clinical signs). Many general toxicity studies, for example, are considered insufficient in terms of the conduct of clinical neurobehavioral observations or the number of samples taken for histopathology of the nervous system. Thus, a battery of negative toxicity studies with these shortcomings would be regarded as providing insufficient evidence of the lack of a neurotoxic effect of the test material. Further. most screening studies based on simple observations involving autonomic and motor function provide insufficient evaluation of many sensory or cognitive functions. Data, which by itself would likely fall in this category, would also include information on structure-activity relationships or data from in vitro tests. While such information would be insufficient by itself to proceed further in the assessment it could be used to support the need for additional testing.

Data from all potentially relevant studies, whether indicative of potential hazard or not, should be included in this characterization. The primary sources of data are human studies and case reports, experimental animal studies, other supporting data, and in vitro and/or structure-activity relationship data. Because a complex interrelationship exists among study design, statistical analysis, and biological significance of the data, a great deal of scientific judgment, based on experience with neurotoxicity data and with the principles of study design and statistical analysis, is required to adequately evaluate the data base on neurotoxicity. In many cases, interaction with scientists in specific disciplines either within or outside the field of neurotoxicology (e.g., epidemiology, statistics) may be appropriate.

The adverse nature of different neurotoxicity end points may be a complex judgment. In general, most neuropathological and many neurobehavioral changes are regarded as adverse. However, there are adverse behavioral effects that may not reflect a direct action on the nervous system. Neurochemical and electrophysiological changes may be regarded as adverse as a function of their known or presumed relation to neuropathological and/or neurobehavioral consequences. In the absence of supportive information, a professional judgment must be made regarding the adversity of such outcomes, considering factors such as the nature, magnitude, and duration of the effects reported. Thus, correlated measures of neurotoxicity strengthen the evidence for a hazard. Correlations between functional and morphological effects, such as the correlation between leg weakness and paralysis and peripheral nerve damage from exposure to tri-ortho-cresyl phosphate, are the most common and striking example of this form of validity. Correlations support a coherent and logical link between behavioral effects and biochemical mechanisms. Replication of a finding also strengthens the evidence for a hazard. Some neurotoxicants cause similar effects across most species. Many chemicals shown to produce neurotoxicity in laboratory animals have similar effects in humans. Some neurologic effects may be considered adverse even if they are small in magnitude, reversible, or the result of indirect mechanisms.

Because of the inherent difficulty in "proving any negative," it is more difficult to document a finding of no apparent adverse effect than a finding of an adverse effect. Neurotoxic effects (and most kinds of toxicity) can be observed at many different levels, so that only a single end point needs to be found to demonstrate a hazard, but many end points need to be examined to demonstrate no effect. For example, to judge that a hazard for neurotoxicity could exist for a given agent, the minimum evidence sufficient would be data on a single adverse end point from a well-conducted study. In contrast, to judge that an agent is unlikely to pose a hazard for neurotoxicity, the minimum evidence would include data from a host of end points that revealed no neurotoxic effects. This may include human data from appropriate studies that could support a conclusion of no evidence of a neurotoxic effect. With respect to clinical signs and symptoms, human exposures can reveal far more about the absence of effects than animal studies, which are confined to the signs examined.

In some cases, it may be that no individual study is judged sufficient to establish a hazard, but the total available data may support such a conclusion. Pharmacokinetic data and structure-activity considerations, data from other toxicity studies, as well as other factors may affect the strength of the evidence in these situations. For example, given that gamma diketones are known to cause motor system neurotoxicity, a marginal data set on a candidate gamma diketone, e.g., 1/10 animals affected, might be more likely to be judged sufficient than equivalent data from a member of a chemical class about which nothing is known.

A judgment that the toxicology data base is sufficient to indicate a potential neurotoxic hazard is not the end of analysis. The circumstances of expression of hazard are essential to describing human hazard potential. Thus, reporting should contain the details of the circumstances under which effects have been observed, e.g., "long-term oral exposures of adult rodents to compound X at levels of roughly 1 mg/kg have been associated with ataxia and peripheral nerve damage."

IV. Dose-Response Analysis

This section describes several approaches (including the LOAEL/ NOAEL and BMD) for determining the reference dose or reference concentration. The NOAEL or BMD/ uncertainty factor approach results in a RfD or RfC, which is an estimate (with uncertainty spanning perhaps an order of magnitude) of a daily exposure to the human population (including sensitive subgroups) that is likely to be without an appreciable risk of deleterious effects during a lifetime.

The dose-response analysis characterization should:

a. Describe how the RfD/RfC was calculated;

b. Discuss the confidence in the estimates;

c. Describe the assumptions or uncertainty factors used; and

d. Discuss the route and level of exposure observed, as compared to expected human exposures. (Specifically, are the available data from the same route of exposure as the expected human exposures? How many orders of magnitude do you need to extrapolate from the observed data to environmental exposures?)

A. LOAEL/NOAEL and Benchmark Dose (BMD) Determination

As indicated earlier, the LOAEL and NOAEL are determined for endpoints that are seen at the lowest dose level (so-called critical effect). Several limitations in the use of the NOAEL have been identified and described (e.g., Barnes and Dourson, 1988; Crump, 1984). For example, the NOAEL is derived from a single end point from a single study (the critical study) and