Peer-Review Draft: Report on Carcinogens Monograph on Trichloroethylene

Supplemental information for Appendix D, Tables D-4 to D-6

(these pages precede page A-61)

July 15, 2014

This Page Intentionally Left Blank

Assessment of potential biases and other study quality characteristics

Each primary study was systematically evaluated for its ability to inform the cancer hazard identification using similar guidelines. In general, studies given the most weight in the evaluation of study quality are those that provide the most valid (i.e., low risk of systematic error or biases) and precise (i.e., low risk of random error) risk estimates, and that have adequate sensitivity (e.g., sufficient power, length of follow-up, and adequate levels of exposure) to detect an effect of exposure. Study strengths and limitations were taken into account in evaluating the reported cancer findings. In addition, studies should accurately report their findings and apply appropriate analytical methods for calculating risk estimates.

Biases in observational studies are often classified into three major categories: (1) selection bias, (2) information bias, and (3) confounding (discussed in Section 3.3.2). Studies with lower potential for bias are generally considered to be the most informative for the cancer hazard evaluation. However, the presence of a potential bias in a study does not necessarily mean that the findings of the study should be disregarded. For example, the effect of confounding may only account for a small percentage of the magnitude of the risk estimate. Therefore, an important step in the process of evaluating biases is to determine the probable impact of the described biases on study results—that is, the magnitude of distortion and the direction in which each bias is likely to affect the outcome of interest (if known). The impact of the potential bias or confounding on the study findings is discussed in the cancer hazard assessment (see Sections 4.1, 5.1, 6.1). In addition, the study's sensitivity to detect an effect of exposure, which is principally a function of study size (specifically, the numbers of TCE-exposed subjects in cohort studies or the numbers of TCE-exposed controls in case-control studies), the length of follow-up, levels of exposure to TCE, as well as the extent of exposure or disease misclassification, should be evaluated. Finally, the adequacy of data (range of exposure) and methods used to evaluate exposure-response relationships should be evaluated.

For this review, the component elements of study quality were evaluated using the questions and guidelines outlined in the protocol (see http://ntp.niehs.nih.gov/NTP/roc/thirteenth/Protocols/TCE_Protocol12-31-13_508.pdf). The guidelines describe the ideal methods and design for each study element. Two reviewers evaluated study quality in concert with input from technical advisors and from a public webinar (http://ntp.niehs.nih.gov/go/tcewebinar).

The study quality elements for each individual study are evaluated and summarized in Tables D-4a (cohort and nested case-control studies), D-5a (kidney and liver cancer case-control studies) and D-6a (NHL case-control studies) and include the following: the potential for selection and attrition bias (unlikely, possible, or probable) and information bias, which includes the quality of the exposure and disease assessment (good, adequate, limited to adequate, and limited), and the likelihood of and concern for disease misclassification (unlikely, possible, and probable) or exposure misclassification. The assessment of exposure misclassification is complex and involves multiple factors such as misclassification that subjects were ever exposed and misclassification of exposure level; thus, labels such as "possible," are not used because they do not capture the complexity of the information. The study sensitivity and exposure-response elements evaluated and summarized in Tables D-4b (cohort and nested case-control studies),

D-5b (kidney and liver cancer case-control studies), and D-6b (NHL case-control studies) consisted of study size and exposure prevalence, reported levels or duration of exposure, and exposure metrics and methods used for the analysis of exposure-response relationships. Section 3 of the monograph discusses these elements across studies.

The terms used for defining the potential for selection or information bias are as follows:

- Unlikely/minimal: Information from study designs and methodologies indicate that they are close to the ideal study characteristics and the potential for bias is unlikely or minimal.
- Possible: Study designs or methodologies are close to but less than ideal, recognizing that in observational studies, there is almost always some methodological or informational limitation and thus some potential for certain types of bias.
- Probable: Study designs or methodologies suggest that the potential for a specific type of bias is likely.

In some cases there is insufficient information to evaluate the level of concern. If adequate information is available, each type of bias is also characterized as to whether it is differential or non-differential. Differential (systematic) biases in the selection of study participants or information assessment are related to both exposure and disease status, and have the potential to bias findings in one direction or another, whereas non-differential (random) biases, which are not related to both exposure and disease, tend to reduce the precision of the risk estimates and often bias the findings toward the null. For example, occupational cohort studies may have limited exposure data across exposure groups, increasing the potential for non-differential exposure misclassification, and may also have the potential for a healthy worker (hire or survival) effect, a type of selection bias that tends to bias findings away from finding an effect (if present) in studies where the comparison group comes from the general population.

Some elements of study quality, e.g., exposure assessment, case ascertainment (Tables D-4a, D-5a, and D-6a) and study sensitivity (Table D-4b, D-5b, and D-6b) were assessed using the terms limited; limited to adequate; and good, based on how close these elements are to an "ideal study" described in the protocol. Briefly these are as follows:

- Exposure assessment: A ranking of good was given to studies having many of the follow elements: industrial hygiene or biomonitoring data, individual detailed job-exposure matrices, job or task descriptions, knowledge of exposure setting, consideration of frequency, confidence and intensity, expert assessment, calendar period-specific exposure data.
- Case-assessment: A ranking of good was given to studies with histologically confirmed incidence data or multiple verified sources of case data, with almost complete follow-up and no evidence of bias in case ascertainment.
- Study sensitivity: A ranking of good was given to studies having many of the following elements: larger numbers of exposed subjects or cases, adequate length of follow-up, high levels of exposure, long exposure duration, large groups or subgroups with a range of exposures from low/medium to high to permit the evaluation of exposure-response relationships, and little concern about exposure misclassification. Factual information on these elements is also presented in these tables.