Recommendations From The "Review of the Hanford Thyroid Disease Study Draft Final Report" by the National Academy of Sciences Committee on an Assessment of Centers for Disease Control and Prevention Radiation Studies From DOE Contractor Sites

HTDS Study Management Team Responses

In December, 2000, the National Academy Press published a report entitled: " Review of the Hanford Thyroid Disease Study Draft Final Report". This 218-page report contained several specific recommendations (summarized in the Executive Summary of the report, pages 17-39). Each of these recommendations is listed below, with a response from the HTDS Study Management Team.

A. EPIDEMIOLOGIC AND CLINICAL METHODS AND DATA COLLECTION

A.1. An adequate review of the cytopathology results is needed.

A.1.a. Request an independent review of the FNA biopsy slides.

Response: The design of the HTDS included interpretation by a cytopathologist of all thyroid cytology specimens obtained by fine needle aspiration (FNA) biopsy of thyroid nodules. These interpretations of thyroid biopsy material were performed by a single, experienced cytopathologist in Seattle during the entire five year clinical evaluation period. During the development of the HTDS protocol, the HTDS investigators considered having a panel of cytopathologists review each set of slides for the HTDS. After considering many factors, it was decided that the consistency of the review offered by one experienced cytopathologist outweighed the possible advantages of using a panel of reviewers. After considerable peer review and approval by a federal advisory panel, this decision was adopted for the HTDS in the early 1990's.

We performed 252 FNA biopsies on individuals with thyroid nodules during the 5 year field work at various clinics throughout Washington State. All of these specimens were reviewed by the same cytopathologist. At the conclusion of the five year clinical evaluation component of the study, there were some peer reviewers who suggested that we have additional quality control assessment performed for the cytopathology results, a suggestion also transmitted by the NAS. We therefore decided to have an independent cytopathological review all 252 sets of slides.

This second review was performed by an experienced cytopathologist from the University of Washington. The results of this review showed some degree of disagreement in 32 cases for an agreement rate of 87.3% (252-32)/252 = 87.3%) for the entire cohort. None of these 32 cases involved a disagreement regarding a diagnosis of thyroid cancer. Detailed evaluation of the reviewer comments in these 32 cases showed that for only 18 cases would there be a possible change in the final diagnosis for a participant, typically involving one type of benign lesion versus another type. Thus the overall agreement with the original study cytopathologist was

92.9% or disagreement of 7.1%. While this result indicates quite good overall agreement, there still was potential disagreement in 18 cases. The HTDS study management team decided that these 18 cases should undergo further review.

A third independent cytopathological review was arranged with two cytopathologists at the University of Washington, who reviewed all 18 cases according to the following protocol:

1. All of the FNA slides for the 18 cases were provided to one of the cytopathologists who reviewed them without knowledge of the original cytology interpretation.
2. A detailed written microscopic description and diagnostic impression was recorded for each case.
3. The original cytology report was then reviewed by the cytopathologist.
4. The reviewer then recorded whether he was in agreement or disagreement with the original opinion.
5. For cases that were in disagreement, the slides were reviewed concurrently with the second cytopathologist (from the University of Washington) to reach a consensus opinion.

The results of this review showed 2 cases of disagreement and 16 cases of agreement between the third reviewer and the original study pathologist. The fourth reviewer concurred with the 2 disagreements of the third reviewer. Neither of the disagreements involved cancer versus benign cytology but rather disagreements between benign versus intermediate probability of follicular neoplasm. In both cases, chronic thyroiditis and hypothyroidism were present. In both cases, the results letter and recommendations that were initially sent to the participants during the study informed the patient that thyroid cancer was very unlikely but could not be ruled out. The results of these additional quality control procedures strengthens the validity of the approach in the original study design of having a single, experienced cytopathologist review all of the FNA specimens.

A.1.b. Request clarification of certain hypocellular specimens with abundant colloid that were classified as benign.

Response: The NAS report (page 54) recommends that additional attention be given to FNA specimens that were classified as benign due to abundant colloid but were acellular or hypocellular. First, hypocellular cases were classified as nondiagnostic unless abundant colloid was demonstrated. If abundant colloid was demonstrated it was thought these were highly likely to be benign lesions in part because the HTDS thyroid examiner team obtained more extensive FNA material than is customary in clinical practice (typically 4-10 aspirations per nodule, providing 10-20 slides of specimen material to the cytopathologist). In addition, during an extensive cytopathology review by 2-3 independent cytopathologists (see above), no case involving abundant colloid with a hypocellular specimen was interpreted to be suspicious or suggestive of malignancy. It is therefore quite unlikely that this category of benign nodules included any thyroid cancer cases. However, it is acknowledged that a hypocellular specimen with abundant colloid cannot entirely rule out a malignancy. While repeat FNA material on these cases is not available, we did perform a separate dose-response analysis excluding all such cases with abundant colloid with hypocellular specimens from the benign category with no change in the results (see Results IX-D.2.c.4).

A.1.c. Clarify the criteria for an adequate biopsy.

Response: The NAS report requests HTDS criteria on adequacy of biopsies in terms of numbers of cells and preparation technique. While there are no clear-cut universal criteria on these issues, it is generally recommended that 6-8 slides of material be available for review. The HTDS study physicians produced on average double this amount and often triple this number of slides for microscopic review. With regard to number of cells, the criteria for adequacy used by the HTDS cytopathologist were consistent with that in clinical practice requiring multiple clusters of thyroid cells (usually greater than 6) on multiple slides.

A.2. The HTDS investigators should indicate for how many potential past thyroid diagnoses they were unable to obtain any medical confirmation, with a breakdown by reported type of thyroid disease and dose.

Response: This NAS recommendation seems to stem from a discussion in the body of the text of the NAS Review (Executive Summary, page 56) which states that "the investigators reported that 37% of the 1264 medical records they sought could not be obtained. (The actual number is 1259 after excluding the requests for an ineligible subject, although a more appropriate number to examine is the 1317 medical records and slides requested). It would be desirable for them to indicate for how many potential thyroid diagnoses they were unable to obtain any medical confirmation, preferably with a breakdown by reported type of thyroid disease". The concern here appears to suggest that for the individuals where medical records could not be obtained, thyroid diagnoses were potentially missed and not included in the HTDS results. However, for all evaluable participants, current diagnostic information at the time of clinic evaluation served as the primary basis for thyroid diagnosis (the basis with highest degree of validity). Thus, even if a medical record could not be obtained, the likelihood of a missed diagnosis is generally low. This is because in most situations the HTDS evaluation will provide valid data regarding whether the diagnosis for which the medical record was sought, is confirmed or not confirmed. For example, in a person reporting a thyroid nodule 30 years ago, the diagnosis will be confirmed by HTDS physicians based on current physical exam and ultrasound scans. An exception would be for a person reporting thyroid cancer, who then had thyroid surgery, and then had missing medical records. However, this occurred in only one individual**.** In addition, the 37% represents the percentage of medical records requested but not received, not the number of participants for whom no medical records requested were obtained. Thus, the 37% figure quoted above by NAS greatly overestimates the possibility that those missing medical records contributed to a missing diagnosis. In order to further clarify this issue, we have added a more detailed description to section V.I.4 and V.I.5 pertaining to the diagnoses for participants with at least one missing medical record or slide.

A.3. The mortality experience should be tabulated in more detail.

Response: The HTDS was not designed to evaluate mortality, and the analyses conducted were never intended to investigate a relationship between Hanford radiation dose and cause of death among those in the cohort who died. Nevertheless, in response to a number of comments and suggestions received after the release of the Draft Final Report, including those from the NAS committee, the mortality analyses were extended to provide additional detail. The analytic methods used and the full set of findings are presented and discussed in considerably more detail in Appendix 23, and in the Discussion (section X.C.1) of the HTDS Final Report.

B. DOSIMETRY

B.1. A single document describing clearly the HEDR dose-assessment methodology, including uncertainties and its implementation by HTDS should be prepared.

Response: The models and computer programs developed by the HEDR Project are described in a series of papers and technical reports published by the HEDR investigators. It is outside the scope of the HTDS to rewrite these, which are the product of a separate research project, into a single document. Additional detail is provided in Section VI of the HTDS Final Report to describe the approach taken in the HTDS to use the computer programs developed by the HEDR Project to estimate individual thyroid radiation dose for study participants.

B.2. Errors in the dose-estimation model should be corrected.

Response: Revisions made to the HEDR model, computer programs, and input data following the NAS review are described in a technical report prepared by investigators from Battelle Pacific Northwest Laboratories, which is included as Appendix 22 in the HTDS Final Report (Napier B, Eslinger P, Ramsdell JV, Hope L. Responses to National Academy of Sciences Review Comments on Dosimetry in the Fred Hutchinson Cancer Research Center's Hanford Thyroid Disease Study. Battelle Pacific Northwest Laboratories, PNWD-3060, 2000.)

B.3. All dose-related uncertainties should be taken into account.

Response: As described in the technical report mentioned above (Napier B, Eslinger P, Ramsdell JV, Hope L. Responses to National Academy of Sciences Review Comments on Dosimetry in the Fred Hutchinson Cancer Research Center's Hanford Thyroid Disease Study. Battelle Pacific Northwest Laboratories, PNWD-3060, 2000; see Appendix 22 of the HTDS Final Report), two modifications were made to the CIDER program regarding uncertainty of dose estimates. The first modification added an option to randomly permute the values of dose conversion factors (DCFs), thereby eliminating an artificial correlation among dose realizations. The second modification allowed the assignment of uncertainty to dietary input values that were specified for participants whose doses were based on CATI data. These two modifications required relatively modest revision of the CIDER code by Battelle investigators, because the original version of CIDER was designed to account for uncertainty of DCFs (which were treated as uncertain for all dose calculations) and dietary intakes (which were treated as uncertain when values are not specified). Accounting for other sources of uncertainty, such as errors in residence histories, sources of food, milk and milk products, and lifestyle specifications would require major revision of the CIDER code, since it contained no provision for uncertainty of these inputs. This was outside the scope of HTDS and was not undertaken as part of the revision of the CIDER program.

Incorporating uncertainties on dietary data reported by CATI respondents had relatively little impact on the dose estimates of the 1979 living evaluable in-area participants whose doses were based on CATI data, or on the estimated uncertainties of those doses. For example, a set of dose estimates for these 1979 participants was calculated using reference diet uncertainties with scale factors SF = 1 (the maximum value allowed in the modified version of CIDER) and $SF_0 = 0.25$. As described in section 7.3 of Appendix 22 of the HTDS Final Report, this applied uncertainties proportional to those of the HEDR default dietary values when consumption values greater than 0

were reported by the CATI respondent. When the CATI respondent reported no consumption, uncertainties of the same relative magnitudes were applied to the 0.25 times the mean of the HEDR default values. These new dose estimates were on average only about 3% larger than the original primary dose estimates that were calculated with no uncertainty applied to dietary input data from CATI. The ratios of the 95th percentile doses to the medians, which provide a measure of the uncertainty of the dose estimates (see section VIII.B.3.a of the Final Report) also increased an average of 3% larger when uncertainty was incorporated in the CATI dietary input data, although the geometric standard deviations (GSDs) increased an average of only 0.8%. In view of the small changes produced by incorporating this additional component of uncertainty, further analyses using these dose estimates were not pursued.

C. STATISTICAL ANALYSES

C.1. A number of key tables should be included in the final report to help readers to interpret the dose-response results.

Response: Tables showing the numbers and percentages of participants with disease and thyroid UDA outcomes have been integrated into the sections for each outcome, rather than gathered into an appendix as they were in the draft Final Report. Note that these tables address the apparent intent of the suggestion to "[expand] the tables of high- versus low-dose results" (page 87 of the NAS Report). The tables presenting "high- versus low-dose results", which are not based on estimated dose categories, provided the kind of "additional set of confirmatory analyses" that the reviewers requested (page 88 of the NAS Report). The dichotomous exposure variable -- relatively high versus low exposure -- is defined in section VIII.B.3.b.2 of the Final Report. Analyses of outcomes in relation to this dichotomous exposure variable, including tables of numbers and percentages of participants with disease and thyroid UDA outcomes, have been integrated into the sections for each outcome, rather than gathered into a single section as in the draft Final Report. A number of other additional tables and figures have been added to the Final Report.

C.2. The HTDS investigators should report on those who were out of the dosimetry area for part of the exposure period and examine the impact of the assumption of zero dose received during such periods.

Response: The NAS Committee was incorrect to state that HTDS assumed that participants received "zero dose" while outside the "dosimetry area", i.e., the approximately 75,000 square mile geographical domain that was defined as part of the HEDR Project (see Figure II.A-1). It is evident from the HEDR results that people living outside the domain could have been exposed to ^{131}I from Hanford. Therefore as noted in section VIII.C.1.a.3 of the Final Report, it was not assumed that participants received "zero dose" while outside that area. As described in Section IX.B of the Final Report, the CIDER program, which was developed by the HEDR Project and used by the HTDS, only calculated thyroid doses received while participants were living within that geographical domain. The boundaries of that domain were defined as part of the HEDR Project based on two considerations: the decreasing reliability of the HEDR model for atmospheric transport and deposition of radionuclides at increasing distances from Hanford, and the likelihood that doses received outside the domain were low. Since the CIDER program did not calculate thyroid doses received while participants lived outside the HEDR geographical domain, only very crude representations of those exposures were possible. Therefore only very

limited scoping analyses of the possible effects of including such exposures were appropriate for the HTDS. The scoping analyses that were performed are described in section VIII.C.1.a.3 of the Final Report, and the results of these analyses are described in the subsections entitled "Scoping Analyses Regarding Out-of-Area Participants" in sections IX.C through IX.Q of the Final Report. These analyses addressed the impact of excluding the 249 living evaluable out-of-area participants (i.e., those who never lived within the HEDR domain between December 1944 and the end of 1957).

The proportion of the 3191 living evaluable in area participants who were outside the HEDR domain varied virtually on a day-by-day basis, but in general increased with the passage of time, reflecting the accumulated effect of participant's families that permanently moved out of the area. Thus on January 1, 1945, a total of 147 living evaluable in area participants were outside the HEDR domain. These accounted for 4.6% of the 3191 living evaluable in area participants (although 967 of these 3191 were born after January 1, 1945). On January 1 of 1946, 1947, and 1948, this increased to 435 (13.6%, with 362 not yet born), 576 (18.1%) and 616 (19.3%), respectively. Due to the complexity of the dosimetry system and its inability to provide anything more than crude approximate dose estimates for scoping analyses, no attempt was made to estimate doses received while the "partial out-of-area" participants were outside the HEDR domain.

If in area participants who had a particular disease or thyroid UDA outcome were more (or less) likely to be in the "partial out-of-area" group, or were more likely to receive higher (or lower) thyroid doses from Hanford's ^{131}I while living outside the HEDR domain, compared to those without the disease or UDA outcome, then CIDER's inability to estimate doses received outside the HEDR domain might bias the estimated dose response for that outcome. In particular such a differential or outcome-related underestimation of actual thyroid doses would tend to reduce an apparent dose response if participants with the outcome tended to receive higher doses while outside the domain. This is because the actual total thyroid doses from Hanford's ^{131}I for participants with the outcome would tend to be underestimated to a greater degree than those of participants without the outcome. However it is very unlikely that decisions made in the 1940s or 1950s to move to a new location which happened to be outside the HEDR domain are related to a child's subsequent development of a disease outcome, nearly all of which were diagnosed much later in life, or to a thyroid UDA observed at the HTDS clinic. Therefore it is unlikely that CIDER's inability to estimate doses received outside the HEDR domain caused HTDS to underestimate the magnitudes of positive dose-responses.

If, as seems much more likely, the magnitudes of the doses received by the "partial out-of-area" participants while outside the HEDR domain were unrelated to disease and thyroid UDA outcomes, then the effect of omitting those doses would be to increase the apparent magnitude of any positive dose-response, since a fixed number of excess cases caused by the actual doses received both in- and out-of-area would be attributed to the smaller doses participants received while inside the HEDR domain. Therefore, CIDER's inability to estimate doses received outside the HEDR domain was likely to cause overestimation of positive dose-responses. Consequently it is unlikely to have caused the HTDS to fail to detect dose-related increases in disease or thyroid UDA outcomes.

C.3. Analyses designed to control for possible confounding by geographic area should be conducted.

Response: The analyses described in the subsections entitled "Effect of Excluding Okanogan and Ferry/Stevens Geostrata" in sections IX.C through IX.P of the Final Report provide more meaningful and directly interpretable results concerning the impact of disease and UDA rate variations between geostrata than the stratified analysis suggested by the Committee. As described in those sections, the slopes of the sex-stratified linear dose-response models are generally increased somewhat by exclusion of the Okanogan and Ferry/Stevens geostrata, as expected since participants in those two geostrata tended to have lower estimated thyroid doses from Hanford's ^{131}I, but somewhat higher outcome rates, compared to the remaining geostrata. However for none of the disease and thyroid UDA outcomes analyzed was a statistically significant dose response observed in these analyses.

C.4. There should be a more thorough set of analyses of thyroid-disease rates by milk-drinking information.

Section IX.A.6 of the HTDS Final Report contains descriptions of dietary consumption levels reported for study participants whose doses were based on CATI data. Section IX.B.3 presents results demonstrating how estimated doses varied in relation to those dietary consumption levels, including consumption of fresh milk and milk products. As expected, since thyroid doses from Hanford's atmospheric releases of ^{131}I depend heavily on location in addition to dietary factors, the correlations between dose and milk consumption are modest (see Table IX.B-12 of the Final Report). Since these correlations were modest, and since individual milk consumption data were available only for the 1979 living evaluable participants whose doses were based on CATI data, no analyses of outcomes in relation to milk consumption were performed. Instead, relatively high and low exposure categories were defined on the basis of both milk consumption and residence history as defined in section VIII.B.3.b.2 of the Final Report (see the subsections entitled "Analysis by Dichotomous Exposure Variable" in sections IX.C through IX.P of the Final Report).

C.5. Confidence intervals should be given and used in the interpretation of the study results.

Response: Confidence intervals have been added throughout sections IX.C to IX.Q of the Final Report.

C.6. The confidence intervals should take into account all the sources of uncertainty in the dose estimates.

Response: Please see the response to the comment that "All dose-related uncertainties should be taken into account" above.

D. STATISTICAL POWER AND INTERPRETATION OF THE STUDY

D.1. The HTDS investigators should describe the sources of uncertainty in as quantitative terms as possible and interpret their results in the light of these uncertainties.

Response: The most useful description of the sources of uncertainty in the dose estimates can be found in the publications and technical reports describing the HEDR dosimetry model. The approaches taken to handling dose uncertainties in the HTDS are described in section VIII.C.2.c of the Final Report, and the results of these analyses are presented in the subsections entitled "Uncertainty" in sections IX.C through IX.Q. The impact of dose uncertainties on the study's statistical power is described in section IX.B.4 and discussed in section X.C.5 of the Final Report.

D.2. The HTDS investigators should recalculate the statistical power of the study, taking into account the dose uncertainties if this proves feasible.

Response: The description of the study's power is provided in section IX.B.4 of the Final Report.

D.3. The compatibility of the HTDS study with other studies of radiation and thyroid disease should be re-examined, taking into account the impact of dose uncertainties.

Response: The comparison of the HTDS results to those of other studies of radiation and thyroid disease is discussed in section X.D of the Final Report.

E. COMMUNICATION OF THE STUDY RESULTS TO THE PUBLIC

E.1. Delivering an unpopular message requires sensitivity to the audience's health concerns and fears. In communications about the HTDS final report, implications for individuals and families that have suffered because of thyroid disease should be carefully explained. If there are plausible alternative interpretations of the results, they should be acknowledged.

E.2. The subcommittee supports CDC's open-communication policy and strongly recommends that it continue. It recommends that a new communication plan be devised for the release of the final HTDS report and accompanying public documents, taking into account the problems that have already been encountered.

E.3. In the HTDS final report and all public documents, any significant changes made from the Draft Final Report should be clearly outlined and explained, and all remaining uncertainties should be noted and explained.

E.4. Careful consideration should be given to how to release controversial reports to the public more effectively. The subcommittee suggests that CDC convene a workshop to discuss this and other communication issues of concern.

Response: The staff of the Radiation Studies Branch of the Centers for Disease Control and Prevention (CDC) will have a summary of responses to these recommendations regarding communication of findings to the public on the CDC web site along with the study report: (http://www.cdc.gov/nceh/radiation/hanford).