Responses to Comments Received from the Public and Members of the Scientific Community.

The HTDS Study Management Team and the CDC received numerous comments from members of the public and the scientific community regarding the HTDS Draft Final Report. Comments were received through the CDC Web site, e-mail, letters and at public meetings. This appendix provides a listing of all of the comments received, along with responses to these comments from the HTDS Study Management Team. Included in the comment list are all comments received by the CDC through the end of the official comment period (July 1, 1999), as well as review comments received from a panel of independent scientists convened by the CDC to review the Draft Final Report. The comments and their responses are grouped according to general topic area. When more than one comment addresses the same question or issue, each such comment is listed and one response is provided. Otherwise, responses are provided for individual comments. Note that page or section numbers mentioned in the comments refer to the Draft Final Report; section numbers mentioned in the responses to comments refer to the Final Report.

I. DOSE ESTIMATION

I.A. HEDR and CIDER

I.A.1 *The HEDR project estimates of the amounts of ^{131}I processed and released in 1959 and 1960 are substantially less than the amounts reported by Warren (1961), which are the source of HEDR release estimates.*

I.A.2 *The HEDR project documents use Warren (1961) as a source but do not evaluate the credibility of his values; for example, he projected an unrealistically high scrubber efficiency.*

I.A.3 *The HEDR project misapplied measured release-factor data from 1959-1960 to the period 1951-1957, when less emission-control equipment was in place.*

I.A.4 *The HEDR project incorrectly accounted for operation of the silver reactors in the B and T plants by inexperienced personnel during the first 18 months after installation in 1951.*

I.A.5 *The HEDR project substantially underestimated the source-term uncertainties for the B, T, and REDOX plants.*

I.A.6 *The HEDR project inadvertently used the medians instead of the arithmetic means of the monthly source terms for the air-concentration and ground-deposition calculations.*

I.A.7 *The HEDR project did not propagate the source-term uncertainties to air concentrations, ground deposition, and doses.*

I.A.8 *The air pathway doesn't account for the topography of the region.*

I.A.9 *The HEDR model doesn't account for changes in rates of release.*

I.A.10 The HEDR model doesn't account for chemical effects of the atmosphere (speciation).

I.A.11 There are other models that are better and that have been tested.

I.A.12 "Researchers didn't know the truth about Hanford's reactor fires and what was actually released."

Response: These comments are directed at the dose estimation system, which was developed by the Hanford Environmental Dose Reconstruction (HEDR) project. The dose estimates calculated using the HEDR system are very useful for the purposes of an epidemiological study such as the HTDS. Therefore, they were used for the primary analyses of the associations between outcomes and exposure, as described in section VIII.C.1. However additional analyses using alternative representations of exposure were also included in order to reduce the study's reliance on HEDR dose estimates (see Section VIII.B.3.b and the subsections entitled "Analysis … in Relation to Alternate Representations of Exposure" in sections IX.C through IX.P of the Final Report).

I.A.13 It is not clear whether the soil deposition estimates were based on meteorological data from 1944-47 or from the 1980s (page 4, Section IV, Study Design).

Response: In the final HEDR model, atmospheric deposition of ^{131}I from Hanford was based on meteorological data from 1944-49 (1). See also the HEDR reports by Ramsdell et al. (2) and Stage et al. (3). No revisions in this regard were made in the Final Report.

I.B Other issues concerning dose estimation

I.B.1 All the animals and crops were also affected. A lot of those crops, particularly alfalfa, were marketed in western Washington State. This alfalfa fed the dairy herds of western Washington and probably contaminated all milk products.

Response: The HEDR system only provides estimates of the thyroid radiation doses people received while living within the 75,000 square mile region in eastern Washington State and adjacent areas of Oregon and Idaho (the "HEDR domain"), shown in Figure II.A-1 of the Final Report. It does not provide estimates of doses that people could have received while living in western Washington State or elsewhere outside this HEDR region. Therefore only limited scoping analyses of the impact of possible exposures received while study participants lived outside the HEDR domain were possible. These scoping analyses are described in section VIII.C.1.a.3 of the Final Report, and the results are described in the subsections entitled "Scoping Analyses Regarding Out-of-Area Participants" in sections IX.C through IX.P of the Final Report.

I.B.2 NTS doses for Stevens, Ferry, Okanogan counties are generally higher than for Benton, Franklin and Adams counties. Perhaps more significantly, the GSD is much higher for Stevens, Ferry, Okanogan counties. It would be useful to include a section that describes how these higher NTS exposures in the HEDR low-dose counties were taken into account and why they were not considered to have constituted a confounding factor.

I.B.3 P. 90 The authors indicated in Section E.3.c. that doses from the Nevada Test Site were calculated for the study subjects. There will be some reviewers who will argue that these doses should have been added to the Hanford doses and incorporated into the final dose response analysis. This issue is addressed in Section VII.D, but it needs to be discussed here, too, or the later section where it is discussed should be referenced here.

Response: In the design and analysis of HTDS, no assumption was made about whether exposure to NTS fallout either was or was not a confounding factor; instead this question was examined in the analysis of the study's results. The rationale for not adding Hanford and NTS doses to produce a total dose is explained in section VIII.D.1 of the Final Report. The calculation of estimated NTS doses, and their analysis as a potential confounding factor or effect modifier, is described in section VIII.D of the Final Report. The results of these analyses are given in the sections entitled "Confounding and Effect Modification" in sections IX.C through IX.P of the Final Report. As described in those sections (and summarized in Section X.C.4), for none of the disease and thyroid UDA outcomes was estimated thyroid dose from the NTS identified as a confounding factor or effect modifier.

I.B.4 Because of including Richland residents (see the previous comment), the HTDS analysis could have been confounded by a significant contribution from the inhalation pathway. It is my understanding that HEDR did not estimate a source term for methyl iodine. It would be useful if the final report included a discussion of this factor.

Response: The HTDS analysis was not confounded by the dose contribution from the inhalation pathway because the inhalation doses are included in the dose estimates calculated by the CIDER program. As illustrated by representative dose calculations in Table 4.4 and Figure 4.18 of the summary HEDR report concerning the atmospheric pathway (4), the estimated inhalation doses were not trivial, ranging up to 21 mGy in 1945 for representative individuals less than 1 year old living immediately east of the Hanford site (up to about 10 mGy for Richland).

The HEDR model assumed that the iodine was in the elemental form at the time of its release during the chemical dissolution processes at Hanford, and was then partitioned between elemental, organic, and particulate forms (5). The organic forms included methyl iodide (CH_3I).

I.B.5 I wonder why someone born at ground zero, like me would fit the same scale as if born at your test area's edge.

Response: It's unclear what is meant by "fitting the same scale" in this comment. However, it should be noted that the HTDS examined the relationship of disease risk and estimated radiation dose to the thyroid. The study was designed to include people with a wide range of doses, from extremely low to the highest doses. Therefore the cohort included, among others, persons who

likely to have lived in early childhood in the downwind counties nearest to Hanford: Benton, Franklin, and Adams. These counties may include the "ground zero" mentioned in the comment.

I.B.6 I don't think the study is accurate because people moved all over.

Response: The fact that people have moved over the decades from the 1940s until the 1990s was dealt with in the design of the HTDS in the following ways.

Definition of the cohort. Many people who were exposed to the ^{131}I from Hanford no longer live near Hanford, or even in the Pacific Northwest. To ensure that the study participants would be as representative as possible of the population of interest, the cohort was defined on the basis of characteristics at birth (birth date and mother's usual place of residence from birth records) without regard to subsequent movement to other locations (section IV.A.1 of the Final Report). Their subsequent movements or "residence history" were then taken into account in estimating their radiation doses, as described below.

Locating and recruiting study participants. The difficulty of locating and recruiting potential study participants was likely to be greater for those who have moved away from the Hanford area. However uniformly extensive efforts were made to trace, locate and recruit every selected person, regardless of current location of residence. This is described in sections V.B and V.C of the Final Report.

Participation in study clinics. To increase the chances that potential participants could attend study clinics regardless of their current place of residence, the study helped to arrange transportation and paid travel costs as described in section V.E.3 of the Final Report.

Estimating radiation doses from Hanford's ^{131}I. In the HTDS, the primary representation of participants' exposures to ^{131}I from Hanford was the estimated thyroid dose calculated using the system developed by the HEDR project. Each participant's individualized dose estimate was calculated using specific information provided by the participant or his/her CATI respondent. The information used to calculate estimated doses included the participant's "residence history", i.e., the locations where he/she lived and the times he/she lived in each location.

Out-of-area participants. The HEDR system only provides estimates of the thyroid radiation doses people received while living within the 75,000 square mile region in eastern Washington State and adjacent areas of Oregon and Idaho, shown in Figure II.A-1 of the Final Report. Some study participants (designated "out-of-area" participants) never lived in this region between mid-December 1945 and the end of 1957, and for such participants, the HEDR system cannot calculate a dose estimate. Since it was quite possible that the out-of-area participants could have been exposed to ^{131}I from Hanford, scoping analyses were performed to assess the impact of their exclusion from the dose-response analyses (see section VIII.C.1.a.3 of the Final Report and the subsections entitled "Scoping Analyses Regarding Out-of-Area Participants" in sections IX.C through IX.P of the Final Report).

Estimating doses from the Nevada Test Site. The release in 1997 of information concerning exposures to radioactive fallout from the NTS provided an opportunity to include some information about these exposures in the analysis of HTDS results. In particular, estimated thyroid doses from NTS fallout were calculated for HTDS participants as described in sections

VI.C and VIII.D of the Final Report. These estimated doses accounted for changes in the participants' places of residence, whether inside or outside the HEDR region.

I.B.7 Many aspects of the study were based on criteria from the Hanford Environmental Dose Reconstruction project which is flawed. For example that study computed the radiation amount from butter using its half life. This is very short sighted of them considering rural eastern Washington (state) farms did not have electricity in the forties. How could home made butter possibly be edible in August of 1945 if kept un-refrigerated for that time period?

Response: The CATI included questions regarding the quantities of raw cow's and goat's milk and milk products (cream, butter, buttermilk, cottage cheese, yogurt, and ice cream combined) consumed by the participant. Thus the CATI respondents were able to report whether or not the participant consumed raw milk products, including butter. These questions were also asked regarding the participant's mother's diet if she was pregnant with or breastfed the participant after December 1944. Of course this kind of detailed information was not available from the Expanded In-Person Interview used for participants without CATI respondents. It should also be noted that additional analyses using alternative representations of exposure were also included in order to reduce the study's reliance on HEDR dose estimates (see section VIII.B.3.b and the subsections entitled "Analysis … in Relation to Alternate Representations of Exposure" in sections IX.C through IX.P of the Final Report).

I.B.8 Suggest doses be given for dichotomous exposure classification. Suggest that categorical analysis should be included in the report.

Response: Two sets of analyses using categorical exposure classifications rather than dose estimates were performed. As described in section VIII.B.3.b of the Final Report, the first of these was based simply on geostrata, and the second on a dichotomous variable accounting for participants' residence histories and milk consumption histories. Distributions of estimated doses are shown by geostratum in Table IX.B-4 and by the dichotomous exposure variable in Table IX.B-13 (section IX.B.3 of the Final Report). Results of analyzing disease and thyroid UDA outcomes in relation to these categorical exposure classifications are described in the sections entitled "Analyses … in Relation to Alternative Representations of Exposure" in sections IX.C through IX.P of the Final Report.

I.B.9 There should be clearer explanation of why age adjustment was done in dichotomous exposure analysis, and not elsewhere.

Response: As described in section VIII.C.2.a.2 of the Final Report, analyses of disease and thyroid UDA outcomes in relation the categorical alternative representations of exposure, i.e., geostratum and the dichotomous exposure variable, were adjusted for a possible effect of age at the time of HTDS examination. This was done because there were small differences in the distributions of age at examination among the geostrata and between the high and low exposure categories (see section IX.A.7 of the Final Report). Note that in the primary analyses of dose responses for disease and thyroid UDA outcomes, the possibility of confounding by age at examination was addressed by the age-adjusted analyses describe in the subsections entitled "Confounding and Effect Modification" in sections IX.C through IX.P of the Final Report.

I.B.10 Did the study consider the possibility that milk, originating near Hanford, was sent to Spokane and therefore consumers in Spokane were at high risk? And "shouldn't this assumption have driven a more exhaustive search in the Spokane region?"

Response: The distribution of commercial milk and milk products throughout the HEDR geographical domain was modeled from a variety of sources of information as part of the HEDR project (6). Estimates of doses received by study participants while living in Spokane (or anywhere else in the domain) took this information about milk distribution into account. Regarding the question about a "more exhaustive search in the Spokane region", we assume this refers to the selection of Benton, Franklin, Adams, and Walla Walla counties as the geostrata for identification of persons likely to have high exposures. As described in Section IV.A.1 of the Final Report, these areas were chosen because the Phase I and final HEDR results indicated that locating people who lived as infants and young children in these counties during the earliest years of Hanford operations provided the best chance of including as many highly exposed participants as practically possible. This was based on the evidence that, while residents of Spokane were indeed exposed to ^{131}I from Hanford, they were likely to have received lower doses than people who lived in the late 1940s in the selected counties.

I.B.11 The HEDR project included both airborne releases of iodine 131 and exposures related to the Columbia River. The HTDS only included exposures to airborne iodine 131. A discussion of this exposure pathway might help the reader assess the importance of omitting this potential exposure from consideration in calculating doses. Additionally, given that you did not include potential exposures associated with the Columbia River, we think that you need to be clearer that you are assessing only exposures related to airborne releases of iodine 131.

I.B.12 HEDR [needs to] to acknowledge that documentation from de-classified documents... ...over the last several years have shown that releases to the river in the '50s were as much as 50% higher than the HEDR study shows.

Response: The HEDR model for exposures related to the Columbia River did not include ^{131}I or any other radioisotopes of iodine (7; 8). Explicit description of the ^{131}I from airborne releases has been added at various places in the Final Report.

I.B.13 (P.87 of draft report) The study uses the information from the study subject about water source to determine whether or not pasture used by the appropriate milk cows was irrigated or not. What is the basis for using this indirect information in this manner? What would have been the impact on dose of misclassifying the cow's source of pasture?

Response: The final HEDR model included several feeding regimes for cattle and goats, to account for the various ways that those animals might be fed as determined by the seasonal availability of various kinds of feed, e.g., fresh pasture grass, green chop, stored feed for cattle. In order to estimate the thyroid radiation dose of a person who consumed raw cow's or goat's milk or milk products, feeding regime(s) had to be specified for the animals in question. The CATI provided an opportunity to try to identify the most appropriate feeding regimes of family cows for participants with CATI respondents (HEDR-specified default feeding regimes were used for participants without CATI dosimetry data). However the CATI was developed and in use for data collection in November 1992, well before the final specifications of the HEDR models for

estimating thyroid radiation doses were known. Therefore, as described in section VI.A.3.a of the Final Report, data that were available from the CATI were used to impute the feeding regime: the main source of water for family cows, and the percentage of feed that was pasture, green chop, or other fresh greens. These two items, along with the location of residence, were used to impute the cow feeding regime. No additional analyses were performed to assess the impact of this issue on dose estimates.

I.B.14 P. 87 The authors state that in April 1996 they brought to the attention of the HEDR Task Completion Working Group and others the issue of inconsistencies in dairies between the HEDR data and the study subjects. Has there been a resolution of this issue at this time? How did HTDS handle this issue in the dose calculations? What impact does this issue have on the dose estimates?

Response: There were two possible reasons for these discrepancies: 1) the HEDR project might have found no evidence of a dairy's operation or supply to the area in question during the period in question (see Deonigi, et al, [9]) for a description of the HEDR commercial milk distribution model); or 2) the CATI respondent may have misidentified a dairy. The majority of these discrepancies were unique, i.e., they occurred only once in all of the CATI interviews performed for HTDS. As noted in section VI.A.3.a of the Final Report there were only 12 instances in which the same discrepancy was found in CATIs of more than one respondent. In eight of these 12, the dairy in question was mentioned by only two CATI respondents. Since the reported discrepancies did not provide definitive evidence of inadequacies in the HEDR commercial milk distribution model, that model was not revised in response to these discrepancies. Therefore HTDS adopted the following approach. Whenever a CATI respondent indicated that the participant consumed milk or milk products from a dairy that did not, according to the HEDR data, serve the area in question during the period in question, the dairy was specified as "unknown" in the scenario file of input data for the participant's dose calculation. This had the effect of assigning the HEDR location- and time-specific default as the source of commercial milk and milk products. If the HEDR model specified that only a single dairy served the location at that time, then that dairy was assumed to be the source of dairy products. If the HEDR model identified two or more dairies that served a region during the time period of interest, then the default was defined as a mixture of milk and milk products from those dairies.

In order to assess the impact of this issue on the dose estimates, it would be necessary to modify the commercial milk distribution model built into the HEDR model. However, because the HEDR project found no evidence that the dairies in question served the specific areas during the time periods in question, no information is available regarding the amounts, and in some cases the sources, of milk and milk products they supplied. Therefore any such modifications would require assumptions about quantities for which there is no direct evidence. Consequently no additional analyses were performed to assess the impact of this issue on dose estimates.

I.B.15 P. 88 The authors state that "…it was impractical to allow a participant's reference diet category to change over time." What was the impact of this computer code limitation on the calculated doses, and thus the dose response analysis?

Response: Reference diet libraries are used by the HEDR model to provide dietary intake values, i.e., quantities of food and milk consumed, when those values are not specified in a participant's scenario file of CIDER input data (see section VI.A.3.a of the Final Report). Since it was impractical to allow participants' reference diets to change over time, it was not possible to assess

directly the impact of this limitation of the estimated doses or dose-response results. However, as described in section IX.B.1 of the Final Report, two alternative sets of dose estimates were calculated. These differed from the primary dose estimates either by replacing the participant-specific information obtained from CATIs with HEDR data (first alternative), or by replacing the HEDR default milk consumption data, when required, defaults derived from the HTDS CATI data (second alternative). The dose-responses for disease and thyroid UDA outcomes were analyzed using the primary and two alternative sets of dose estimates. As shown in the subsections entitled "Analysis … in Relation to Alternative Dose Estimates" in sections IX.C through IX.M, IX.O, and IX.P of the Final Report, for neither alternative set of doses were the study's findings changed. In addition, the source of dosimetry data (CATI versus Expanded In-Person Interview) was treated as a potential confounding or effect modifying factor. Since the CATIs provided specific dietary intake values for most participants with CATI data, while no specific dietary data were obtained from the Expanded In-Person Interviews, this analysis reflected in part the effect of the selection of the backyard cow's milk reference diet library. No further analyses were conducted to investigate the impact of the limitation to a single reference diet.

I.B.16 What is the impact on the dose estimates of using "fuzzy date codes" for residences, and how did you evaluate this impact?

Response: The "fuzzy date codes" were a set of conventions that were used when CATI respondents or study participants were unable to provide exact dates for events such as changes in residence or dietary practices. For example, if an interviewee could only specify the month within which an event occurred, the event was assumed to have occurred on the 15th of that month. If only the year was known, then the event was assumed to have occurred on July 1 of that year. The following dates were used if only the season of the year was known: February 1 for Winter, May 1 for Spring, August 1 for Summer, and November 1 for Autumn.

The CIDER program required specification of exact dates in each participant's scenario file, making it impractical to investigate the impact of uncertainties in event dates. Therefore the potential impact of this on the estimated dose responses was not investigated.

I.B.17 Consistency among individuals' dose estimates is lost when each individual's dose is represented by the median estimate.

I.B.18 The final report should include a discussion that addresses why the loss of correlation (among the CIDER realizations) was not considered to be a problem when most of the dose-response analyses relied on only one value of each participant's dose estimates (i.e., the median).

I.B.19 Question the loss of correlation because each person's median dose comes from a different [could not read].

I.B.20 Is it true that higher and lower estimated doses were actually higher and lower median doses? And that most of the dose response analysis in the study is based upon median values?

Response: The primary descriptive and inferential analyses using individual dose estimates were based on the median of each living evaluable in-area participant's 100 dose realizations

calculated by the CIDER program; see section VIII.B.3.a of the Final Report. The statement in the first comment above is true, and it is also true for the two other point estimates of dose in section VIII.B.3.a of the Final Report, the geometric and arithmetic mean doses. However the practical value of having a single point estimate of dose for both descriptive purposes and analyses of the magnitude of exposure and dose-response relationships far outweighs this criticism. The second comment is incorrect in its assumption that "loss of correlation" arising from the use of medians was not considered a "problem". As just mentioned, estimated dose-responses based on the medians are of immense practical value. In addition, sections IX.C through IX.M, and IX.O and IX.P provide results estimates of the slope of the dose-response parameter for each of the sets of 100 realizations produced by the CIDER model (see the figures entitled "Plot of Estimated Slope by Dose Realization).

I.B.21 If 2/3 of the people had some type of thyroid malfunction then [speaker] would have liked to see a birth date quantity chart in the report, so he'd know who was born when, and a map that could easily tell where the doses fell.

Response: Birth years of the 3440 living evaluable study participants are tabulated in Table IX.A-1 in section IX.A of the Final Report. Maps illustrating the areas in which participants were likely to receive comparatively high or low exposures have been published by the HEDR project (7, 4). An example of such a map is shown in Figure IV.A-1 of section IV.A in the Final Report. However, it must be understood that such maps, while useful for descriptive purposes, do not accurately reflect the doses actually received by all study participants, since they do not account for participants' individual residence histories and dietary histories. The dose estimates used in the HTDS used such individual information to the extent it was available. Average estimated thyroid doses from Hanford ^{131}I are shown by sex, birth year, and geostratum in Figure IX.B-5 in section IX.B.3 of the Final Report. The cumulative incidence of disease outcomes and prevalence of thyroid UDAs is shown by sex and geostratum in the sections entitled "Analysis by Geostratum" in sections IX.C through IX.P of the Final Report.

II. STATISTICS

II.A. Statistical power

II.A.1 Was there sufficient statistical power?

II.A.2 What was the statistical power?

II.A.3 Because of the large number of participants, HTDS achieved a high level of statistical power. The final report should describe the statistical analysis and the confidence levels for the conclusions based on the least squares analysis.[based on summary]

II.A.4 I have heard from colleagues that a number of scientists consider HTDS to be inconclusive because of its low statistical power.

II.A.5 (the study was released) with the comment that the results were powerful, but other analysis suggests it is not as high.

Response: Section IX.B.4 of the Final Report contains a much more comprehensive discussion of the study's power than was provided in the draft Final Report.

II.A.6 Statistical power should be expressed as a subjective probability distribution.

Response: The HTDS investigators disagree with this approach to expressing statistical power. See section IX.B.4 of the Final Report for description of the study's statistical power.

II.A.7 How sensitive is the computed power to the targeted excess probability for disease?

Response: The relationship of the study's statistical power to the hypothesized value of the excess probability of disease or thyroid UDA outcomes is illustrated in Figures IX.B-7, IX.B-8, and IX.B-9 in section IX.B.4 of the Final Report.

II.A.8 How sensitive is the computed power to overestimation of dose by factor of three?

Response: The impact of dose uncertainties on the study's statistical power is described in section IX.B.4 of the Final Report.

II.A.9 How sensitive is the computed power to underestimation of the average background probability of disease?

Response: The impact of background rates on statistical power can be seen by comparing the results for outcomes with relatively low, intermediate, and high background rates, as described in section IX.B.4 of the Final Report.

II.A.10 On page 3 of the Introduction, you state that you had sufficient power to detect an increase of 5% in thyroid neoplasia per Gray. A clarification of this point would help us better understand what you mean by "relatively small".

Response: The subjective term "relatively small" has been deleted from the Final Report. Section IX.B.4 includes a description of the context within which to interpret the "detectable" effects.

II.A.11 (question about) the difference in interpretation of statistical power of the study between the FHCRC and CDC and why this has not been brought out publicly….

Response: Because this comment originated at a public meeting and the exact text is not available, it is not completely clear what this refers to. The implication is that the questioner was under the impression that the CDC staff disagreed with the HTDS investigators regarding the interpretation of the power of the study. There was extensive discussion between the CDC and the HTDS investigators prior to the release of the draft report regarding the interpretation of the

primary findings of the study, the power of the study, and the fact that the effects of dose uncertainty were not yet incorporated into a quantitative estimate of study power, or the quantitative estimates of risk from the dose-response analyses. Within this context the CDC staff reviewed and approved of the release of the summary findings booklet, the study fact sheets, and the prepared statement that was delivered at the press conference. There was agreement between CDC staff and HTDS investigators regarding the key messages from the study and the primary conclusions.

II.B. Uncertainty

II.B.1 The study's dose-response analysis is incomplete because the dose uncertainty has not been fully incorporated, as called for in the HTDS final analysis plan.

Response: The methods for analyses accounting for dose uncertainties are described in section VIII.C.2.c of the Final Report. Results of these analyses are included in the subsections entitled "Uncertainty" in sections IX.C through IX.Q of the Final Report.

II.B.2 What does uncertainty in dose estimates do the computed power?

Response: This is discussed in section IX.B.4 of the Final Report.

II.B.3 Maximum likelihood estimates are only one possible set of parameter values of the risk model. In fact, they may often be a next to arbitrary choice from a set of possibly applicable parameter values.

Response: Maximum likelihood estimation is a standard, widely used, and well understood technique for estimating parameters of statistical models, and is quite appropriate for the HTDS. Nevertheless, in response to the recommendation of the NRC Committee, estimates were also calculated by an alternative method, the method of least squares; see section VIII.C.2.a.4 of the Final Report. Results of both maximum likelihood and least squares analyses of dose-responses for disease outcomes and thyroid UDAs are presented in sections IX.C through IX.P of the Final Report.

II.B.4 An explanation of the sensitivity of the model to different levels of reporting of dietary intake might also help the reader in assessing the importance of misclassification due to inaccuracies in reporting of diet.

Response: The meaning of "levels of reporting of dietary intake" in this comment is unclear. The two possible sources of dietary data for study participants were the CATI, which sought detailed individual dietary histories, and the Expanded In-Person Interview, which collected no individual dietary history other than sources of cow and/or goat milk. As described in section VIII.A.1.b of the Final Report, the source of dosimetry data was investigated as a possible confounding or effect modifying factor. Results of these analyses are given in sections IX.D through IX.M, IX.O and IX.P in the tables entitled "Confounding and Effect Modification by Sex, Age at Exposure or HTDS Examination, Estimated Thyroid Dose from NTS, History of Any Cancer Other Than Thyroid and Interview Type". For none of the outcomes analyzed was there

evidence that estimation of the dose-response was significantly confounded or modified by differences in the source of dosimetry data.

II.C. Statistical (other)

II.C.1 There is concern about the characterization of our confidence in the results of the specific outcomes.

Response: Confidence intervals have been added throughout sections IX.C to IX.Q of the Final Report. These provide a description of the precision with which dose-response parameters were estimated.

II.C.2 The p-values for most analyses are very high. This signifies very little confidence in the results.

Response: This assertion in this comment that "very high p-values" signify "very little confidence in the results" is incorrect. Since the dose-response analyses of disease and thyroid UDA outcomes emphasize the one-sided alternative hypotheses that risk increases with dose, whenever the estimate of the dose-response parameter (e.g., the slope B in the linear probability model) has value less than 0, the corresponding p-value will be greater than 0.50.

II.C.3 Many of the upper bounds of 95% confidence intervals are in the positive side (implying a positive dose-response).

Response: The upper confidence limits are positive because the estimated dose response parameters (i.e., the "point estimates" of the slope B in the linear probability model) are close to zero. Confidence intervals for the dose response parameter range from a lower limit less than the point estimate to an upper limit greater than the point estimate. Therefore, for example, if the point estimate of a slope is exactly zero, the upper confidence limit will always be positive, i.e., greater than 0. Therefore the fact that the upper confidence limits are greater than 0 does not provide convincing evidence that disease risk increased with increasing dose.

II.C.4 Considering the HTDS use of surrogate dosimetries, there is an error in how the second alternative representation of exposure is described in the draft report. In section C.3 of the Discussion (IX). This was confirmed via a telephone conversation with Ken Kopecky on February 26, 1999. The final report should include an accurate description of how this analysis was performed.

Response: The definition of the high and low exposure categories has been revised from the version described in the draft Final Report. In particular, more detailed information regarding how participants' milk consumption histories were used (see section VIII.B.3.b.2 of the Final Report).

II.C.5 In doing the geostratum surrogate, was it adjusted for residence location in 1945? If not, this should be considered for inclusion in the final report.

Response: The analyses of outcomes in relation to geostratum were not adjusted for residence in 1945. Rather than perform such an adjusted analysis of differences between geostrata, a more straightforward approach would be to simply analyze outcomes in relation to residence in 1945, if the latter can be defined in a meaningful and defensible way. Such an analysis, if possible, would likely address the intent of this comment. However it is not clear how, for either kind of analysis, "residence location in 1945" should be defined for persons who lived in multiple locations in 1945. Should it be the first residence location in 1945, location of longest residence during 1945, location of residence during period of highest ^{131}I releases from Hanford, something else? All of these possible definitions are subject to the criticism that they may misclassify some participants with respect to the relative sizes of their thyroid radiation doses from Hanford's ^{131}I.

The analyses of outcomes in relation to a dichotomous exposure variable (see section VIII.B.3.b.2 of the Final Report) actually goes a long way to meeting the concern expressed in this comment. For the Final Report, this analysis has been enhanced by including both residence history and milk consumption history in the definitions of the relatively high and low exposure groups.

II.C.6 There should be discussion of mis-classification of outcomes, specifically noting results of ultrasonography QC program on page 108 of procedures section.

Response: Misclassification of outcomes is discussed in section X.C.2 of the Final Report, and the ultrasonography quality control studies are described in section V.F.9 of the Final Report. The HTDS procedures for clinical evaluation, interviews, medical record collection, and review and final diagnostic determination, described in sections V.D and V.F through V.I of the Final Report, were designed to provide highly reliable information about the presence of thyroid disease in the study participants. The primary analyses of disease outcomes were based on the most definitive diagnoses, based on a comprehensive review of all available pertinent information in the final diagnostic determination (section V.H). In addition, broader but less definitive alternative definitions were established for each disease outcome in order to investigate whether dose-response results might be influenced by the inclusion of less well-documented diagnoses (see the subsections entitled "Alternative Definitions for Diagnosis …" in sections IX.C, IX.D, and IX.F through IX.M, and IX.O of the Final Report). Any misclassification of outcomes that occurred was unlikely to be a source of bias in the estimation of dose-responses. This is because care was taken to ensure that participants did not reveal information about their possible exposure levels to the physicians and ultrasonographers at the study clinics (see sections V.F.2.d and X.C.2 of the Final Report). Also the procedures for collecting and reviewing outcome information were designed to ensure that the final diagnostic determination was made without knowledge of factors that might influence the participant's thyroid dose from Hanford's ^{131}I.

II.C.7 The sonographers had somewhat greater differences than the radiologists including discrepancies in number of nodules and presence or absence of nodules greater than 5mm. These discrepancies are not addressed and should be commented on briefly.

Response: There is little reason to compare the levels of agreement observed among radiologists to those observed among sonographers. This is because the radiologists' reviews were performed under quite different circumstances than the sonographers' examinations. As described in Section V.F.9 of the Final Report, the comparisons among radiologists were based on their reviews of videotape records of sonographers' examinations, while sonographers were compared

on the basis of the results of actual examinations that they separately performed. Therefore the radiologists saw exactly the same images, while the sonographers saw different images, as determined by their individual examining techniques.

II.C.8 The report needs more extensive discussion of multiple statistical comparisons.

II.C.9 On page 90 of section VIII and on page 13 of section IX, mention is made of conducting a large number of significance tests in the context of secondary and alternative analyses and the need for caution in interpreting a specific p-value of 0.003. This caution applies study-wide, not just to this test alone. Can the few significant results observed be explained by chance alone?

Response: The problem associated with performing multiple comparisons, i.e., the likelihood that some comparisons will be falsely significant due to chance alone, increases the risk of "false positive" results. The impact of this problem on the interpretation of dose-responses is discussed in Section X.C.3 for nonpalpable focal thyroid UDAs and for diffuse UDAs, the two disease or UDA outcomes with nominally statistically significant dose-responses (at least in the analyses that excluded participants with doses over 400 mGy).

The calculation of multiple confidence intervals is subject to a related problem. Specifically, if a confidence interval for a parameter is calculated at a given confidence level, e.g., 95%, then, loosely speaking, one can be 95% confident that the true parameter value lies within the interval. However if multiple confidence intervals are calculated, each at the given confidence level, then the overall level of confidence that all of the true parameter values lie within their respective intervals is less than the nominal level. Consider, for example, the three parameters of the sex-stratified linear probability model [1] in section VIII.C.1.a of the Final Report: one can be no more than 86% confident that all three true parameter values lie within their respective 95% confidence intervals. The approach taken to address this problem, called the Bonferroni technique, is described in section VIII.C.2.b.4 of the Final Report.

II.C.10 The words used to describe levels of statistical significance in the analysis are sometimes imprecise or contradictory. It may be inferred from the document that the authors intend to define marginal statistical significance as $0.01 <= p < 0.05$ and statistical significance as $p < 0.01$. This needs to be explicitly stated and whether or not these definitions were made a priori.

Response: Adjectives characterizing statistical significance have been omitted in the Final Report.

II.C.11 There was inconsistent applications of effect modification.

Response: The results of analyses of effect modification are given in detail for disease and thyroid UDA outcomes in sections entitled "Confounding and Effect Modification" in sections IX.C through IX.P of the Final Report. These analyses were not performed for the outcomes of thyroid cancer, other thyroid disease, or Hyperparathyroidism due to the small numbers of cases.

II.C.12 Statistical models should be specified more clearly. It was difficult to tell when models were age-, sex-adjusted and when they were not.

Response: Sex-stratified linear or logistic regression models were used for all analyses of disease and thyroid UDA outcomes in relation to estimated thyroid radiation dose from Hanford's ^{131}I or to alternative representations of exposure, as described in sections VIII-C.1 and VIII-C.2 of the Final Report. Note that this includes analyses of potential confounding and effect modification, and analyses accounting for dose uncertainties.

Analyses were generally not age-adjusted, since the cohort was defined to focus on persons with a fairly narrow range of age at first exposure to Hanford's ^{131}I, who consequently also had a relatively narrow range of age at HTDS examination (see Table IX.A-1 in section IX.A of the Final Report). Note however that age at first exposure to Hanford's ^{131}I and age at HTDS examination were included as potential confounders and effect modifiers; see section VIII.A.1.b and the subsections entitled "Confounding and Effect Modification" in sections IX.C through IX.P of the Final Report. Also, analyses of disease and thyroid UDA outcomes in relation to categorical alternative representations of exposure (geostrata or a dichotomous [high versus low] exposure variable) were adjusted for age at HTDS examination; see section VIII.A.3.b.

II.C.13 … in 1990 with preliminary results of the HEDR project, the doses looked fairly large. ..in 1994 the dosimetry was refined and the dose levels were much lower. …power calculations were borderline of whether or not the study could go forward... [he asked] why the complete analysis plan had yet to be carried out and implemented.

Response: The considerations and review leading to the decision to proceed with the Full Study are described in section V.A of the Final Report.

As described at the time the draft Final Report was released, complete results concerning the effect of dose uncertainties on estimated dose-response relationships were not available because the method of analysis proposed in the HTDS analysis plan had proven impractical. A revised approach for these analyses was undertaken for the Final Report. This revised approach is described in section VIII.C.2.c of the Final Report, and the results are presented in the subsections entitled "Uncertainty" in sections IX.C through IX.Q of the Final Report.

II.C.14 Question about whether or not by using the dose information from all the study participants, both those with CATI derived dose estimates and those with HEDR default derived dose estimates, whether or not that obscures the real dose response analysis because of the upwards bias of using the HEDR default for some of the doses and not for all.

Response: This was addressed in analyses that treated the source of dosimetry data (CATI versus Expanded In-Person Interview) as a potential confounding factor or effect modifier. The methods for these analyses are described in section VIII.C.1.b of the Final Report, and results of these analyses are described in the subsections entitled "Confounding and Effect Modification" in sections IX-C through IX-P.

II.C.15 It may be desirable to do another dose response analysis using the number (or size and number), and not just the presence of UDAs. A dose response for number of lesions might be possible for larger lesions, e.g. greater than 10 mm, but not for greater than 5 mm lesions.

II.C.16 The study should consider if a dose response analysis of the number and size of UDAs is feasible and should be performed (taking into account my caveat regarding sonographer concordance).

Response: The suggested analyses can be found in section IX.P.2.k of the Final Report (see also sections VIII.C.1.d and VIII.C.2.a.3).

II.C.17 .. it is important to realize the difference between a dose study and the link issue. paraphrased: if the study doesn't find a dose response it doesn't mean there isn't a link.... that is not the way the study is set up ... .

Response: The HTDS was basically a typical epidemiological study, similar in design and execution to a very large number of other studies that have been conducted over the years to investigate the potential health impacts of exposures to a wide variety of potentially harmful agents, including ionizing radiation. The methods that have been developed for these studies take into account the fact that individual cases of disease, or of other outcomes such as thyroid UDAs, cannot be directly linked or attributed to a particular exposure. This is because (1) the outcomes being studied, thyroid and parathyroid disease and thyroid UDAs in the case of the HTDS, can all occur spontaneously, i.e., in the absence of exposure to ^{131}I, and (2) there are currently no known markers or other characteristics that distinguish cases of disease caused by radiation exposure from the spontaneous ("background") cases. Since it is not possible to detect direct, causal links between exposure and outcomes, an epidemiological study is designed to search for statistical evidence of associations between exposure and disease risk. The statistical results of such a study, while they cannot provide absolute proof of the presence or absence of causal links between exposure and outcome, nevertheless provide information of great importance to the public, the medical community, health officials, and scientists. The inappropriateness of requiring absolute proof of the presence or absence of a causal link, and ignoring statistical evidence for the presence or absence of exposure-outcome association, is perhaps best illustrated by the example of smoking. The tobacco industry argued for decades that the harmful effects of smoking were not "proven" because the evidence for those effects, as overwhelmingly compelling as it is, came largely from epidemiological studies that showed statistical associations between smoking and disease.

II.C.18. The report should emphasize prominently the difference between the primary analyses, formulated as part of the study design, and secondary analyses, conducted after the data have been obtained. [Suggest this is moved into Statistics somewhere]

Response: In sections IX.C through IX.M, IX.O, and IX.P, the primary analyses (i.e., those based on the primary definition of the disease outcome and the sex-stratified linear model) are placed in subsections entitled "Primary Analysis." These are subsections 2.a in IX.C through IX.M and IX.O, and 2.a, 3.a, 4.a, and 5.a in IX.P. As noted in section IX.N.1, the primary analysis was not performed for the outcome of Other Thyroid Disease since there were too few cases for meaningful analysis.

II.C.19 Why not use confidence interval instead of standard error

Response: Confidence intervals, calculated using the Bonferroni method to adjust for the simultaneous estimation of multiple parameters, have been added to dose-response results throughout chapter IX of the Final Report (see also section VIII.C.2.b.4).

III. DOSE RESPONSE

III.1 Chronic low dose radiation is more effective in producing health effects

Response: The assertion in this comment is not supported by the medical and epidemiological literature. It has been well substantiated that acute, high dose radiation exposure is more harmful and associated with numerous short term and long term health effects, than is chronic, low dose radiation. See section II.B of the Final Report for a review of the current literature regarding radiation and thyroid disease risk.

III.2 A least squares analysis would show the actual relationships between the dose and the effects. It is obvious from analysis of Figures 1 and 2 that a least squares analysis would show that the frequency of cancer decreases with dose.

Response: The method of least squares is an alternative to the method of maximum likelihood for estimating the parameters of the linear dose response models. These two methods can be expected to give similar though not identical results for disease and thyroid UDA outcomes. Although it adds little to the analysis and has no impact on the study's findings, least squares estimation of linear dose response models for disease and thyroid UDA outcomes has been added. The application of least squares is described in section VIII.C.2.a.4 of the Final Report, and the results of least squares estimation are described in the subsections entitled "Primary Analysis" in sections IX.C through IX.P of the Final Report.

IV. CONTROL POPULATION

IV.1 Why wasn't there a control population?

IV.2 I think a wider study is called for and a control area is needed that is outside the reach of the farmers market.

IV.3 The decision to use a low-dose rather than no-dose comparison group may also limit the study's ability to detect effects due to exposure to iodine 131.

Response: As explained in section IV.A of the Final Report, the HTDS adopted the approach of using one population comprised of individuals with different levels of exposure to radiation rather than two separate populations (one exposed and one unexposed) to see if there was a relationship between exposure to Hanford radiation and the risk of thyroid disease. This approach has been used extensively in assessing the effects of radiation exposure in human populations. It is a common design in epidemiology, and has been employed in studies of atomic bomb survivors in Japan, in numerous studies of people exposed to radiation through medical procedures, and in the

study of people exposed to radiation from atmospheric testing in Utah. This method is superior to the alternative approach of attempting to compare thyroid disease occurrence in a cohort under extensive study such as the HTDS cohort with that in a separate population presumed to be unexposed to radiation. This is because thyroid disease rates may be a function of a number of factors other than exposure to radiation. These factors may differ considerably between different populations, particularly if one population is under careful study and diagnostic evaluation. Such differences can include: 1) the methods of diagnosis employed; 2) the extent to which diagnostic tests are implemented in a population (i.e., the thoroughness of the diagnostic process); 3) the dietary practices of the population; 4) the level of stable iodine in the diet; and 5) the composition of the population according to age, gender, and ethnicity. To the extent differences in such factors exist, it would be impossible to attribute any differences in thyroid disease rates observed to Hanford radiation exposure, as opposed to one or more of these other factors. The approach used in the HTDS is also superior to one that would implement the full HTDS protocol in a population geographically removed from the Hanford, in an attempt to include persons with no exposure from Hanford radiation. Although the methods and thoroughness of the diagnostic evaluation would be comparable under such circumstances, it would still not be possible to ensure comparability between the two study populations regarding the other types of possible differences listed above that could influence thyroid disease occurrence. Thus, to ensure as much comparability as possible regarding factors other than radiation that can influence the occurrence of thyroid disease, all comparisons of thyroid disease rates in relation to thyroid radiation dose level were made within the defined cohort.

V. STUDY DESIGN AND SELECTION CRITERIA

V.1 *To exclude (name) [born outside of study area] from any consideration that her health might have been injured due to Hanford emissions is an affront to most intelligent people. We believe that this study did not involve a large enough sampling nor did it cover enough years. Why did it stop at 1946? Was it assumed that nobody born beyond 1946 could have been affected?*

V.2 *I would like to know why the study was done on people born in 1940-46 when people were exposed during 1944-1957?*

V.3 *Says study was not "fair" since not everyone with thyroid disease was included.*

V.4 *Did the study include the "right" counties? Didn't we miss legitimate counties?*

V.5 *Why is Okanagon County a low dose area?*

V.6 *Why wasn't the higher population from Spokane added?*

V.7 *Why weren't the migrant worker population and Native Americans in the study?*

V.8 *Even though I was born in Benton County in 1941, we moved to San Diego a couple years later. Except for an occasional visit to grandparents, I did not spend much time in that geographical area. And, Benton County is not normally 'downwind' from Hanford anyway. Which brings the sampling into question in my opinion.*

V.9 *My ex-wife [who had Hashimoto's] was born in southern Idaho, but her family moved to Cheney, Washington, when she was very young and she had lived there until adulthood. The study people refused to accept this information since she was not born in the counties they were sampling. This told me that they just didn't want to hear the truth, and that a vast number of "Downwinders" would not have a chance to be a part of the study.*

V.10 *How a bunch of so-called "experts in their field" can make the sweeping statement that there is no cause-and-effect between the Hanford emissions and thyroid disease is beyond my understanding. Nobody studied [my wife] or the year she was born in. How can [my wife's] case be so cavalierly dismissed when she was never studied or examined as an individual?*

V.11 *Were they even aware that a portion of Grant County is within the boundaries of the Hanford Nuclear Reservation?*

V.12 *I am dubious of the study's age group and geographical mix.*

V.13 *Spokane's directly NE-Downwind- from Hanford. Why wasn't sampling done from a larger, concentrated population?*

V.14 *It is incomprehensible that the Hanford thyroid study did not include 25,000 students at Washington State University for that period of time.*

Response: It is not possible to include in a study like the HTDS everyone who was ever potentially exposed to Hanford radiation. More importantly, it is not necessary to do so in order to achieve the primary objective of determining whether Hanford radiation exposure resulted in an increased occurrence of thyroid disease. If the study is conducted correctly, the results will be meaningful to a much broader population of people than just those relatively few who actually were in the study. The most important principle to follow in conducting a study like this one is to select people for study in an unbiased manner; that is, *not* based on the knowledge that they lived around Hanford and developed thyroid disease, or that they didn't live around Hanford and didn't develop thyroid disease. Such a design is "fair" because it includes people who were potentially exposed to Hanford's ^{131}I, without regard to whether or not they developed thyroid disease. It was critical that we identified a group of people exposed to Hanford radiation, and then determined in an unbiased way exactly what happened to them thereafter regarding the development of thyroid disease. The most important considerations were to include people who were exposed to Hanford radiation, to include persons with the highest exposures as well as those with little or no exposure, to select people without knowledge of whether they have thyroid disease, and to collect information from every study participant in exactly the same way regarding their exposure and their thyroid disease status.

Because we did not know at the start of the study which individuals were exposed to Hanford radiation or how much exposure they might have received, for the purposes of subject selection only, residence at time of birth acted as a surrogate for the anticipated radiation dose to the thyroid from Hanford. Individual thyroid radiation dose could only be estimated from data collected during the study. As noted in section IV.A of the Final Report, preliminary findings from the HEDR project at the time the study began regarding meteorological conditions affecting the deposition and concentration of radioactive iodine in vegetation, and the patterns of milk production and consumption by county, indicated that persons with the highest thyroid doses were most likely to have lived in the area encompassed by Benton, Franklin, and Walla Walla

counties. Thus, these counties were targeted as areas where we might identify persons with the highest doses from Hanford radiation. The selection of cohort members was also extended to include three counties on the Canadian border north of the Hanford site (Okanogan, Ferry and Stevens). These counties were selected because, based upon the information available at the time regarding possible radiation doses to the thyroid, they could be expected to contribute some individuals with very low radiation doses to the thyroid from Hanford. In addition, persons living in these counties would likely be comparable to the group of those who did receive a thyroid dose in terms of other factors which could potentially influence the risk of thyroid disease (e.g., geography, urban/rural composition, occupational factors, socioeconomic factors, age, ethnicity, sex). It was also important that similar opportunities and resources existed to identify and trace persons in these counties as there were in the group that lived in counties closer to Hanford.

Preliminary estimates from the HEDR project also suggested that the highest thyroid doses were most likely to be in persons exposed as infants or children during the first years of Hanford operations. This is because infants and children receive higher thyroid doses per unit exposure due primarily to the small size of their thyroid glands, and existing literature suggests that radiation-induced thyroid disease (and possibly hyperparathyroidism) is greatest among those exposed at youngest ages. For this reason, the study focused on persons who would have been young children at the time that the majority of releases of radioactive iodine from the Hanford facility occurred (1944-46). The best way to identify people who would have been young children during this time in an unbiased manner was to use a roster of all births that occurred around that time period in the counties of interest. We selected persons born from 1940-46 to achieve this end, which meant that the cohort would contain persons with exposure beginning as early as the prenatal period, and as late as age three. We did not purposefully exclude any particular group (e.g., Native Americans). This is not to imply that persons born after that time were not exposed, or potentially affected. This approach was taken to focus on those who most likely received the highest exposure, and who were likely to be most sensitive to the effects of that exposure. An additional benefit of choosing this young age group was that mothers and close relatives of persons born from 1940-46 would more likely be alive and available for interview compared to those of persons born earlier.

V.16 How do we know that Hanford emissions, as well as emissions from other nuclear projects and testing, did not contaminate the entire food chain nationwide, thereby causing thyroid disease on a large scale? If this were the case, then of course a select few Hanford downwinders thyroid problems would not stick out, or be obvious.

Response: The results of the Hanford Environmental Dose Reconstruction (HEDR) Project strongly suggest that the doses caused by ^{131}I released into the atmosphere from Hanford were highest in people who lived in the counties immediately to the east and northeast of Hanford (see, e.g., Figure IV.A-1 in section IV.A of the Final Report) during the first several years of operations at Hanford, i.e., the years of highest ^{131}I releases. This does not mean that people who lived in other areas were not exposed to Hanford's ^{131}I. Indeed results of the HEDR Project indicate that people who lived outside the region shown in Figure IV.A-1 during 1945-57 were probably exposed to Hanford's ^{131}I. However the further away from Hanford a person lived, the lower his or her thyroid radiation dose from Hanford's ^{131}I is likely to be. This is important because studies of the health effects of radiation exposure consistently show that those effects are dose-dependent. That is, the risk of having a health effect increases with increasing dose. Therefore, as described in section IV.A of the Final Report, the study was designed to include as many of the most highly exposed people as possible, since they would be most likely to have

suffered health effects from Hanford's ^{131}I. In addition, the thyroid doses from Hanford's ^{131}I were expected to be larger than those from atmospheric fallout from the Nevada Test Site or other nuclear weapons tests around the globe for many study participants. Therefore an appropriate analysis of the dose-responses for thyroid health outcomes in relation to estimated dose from Hanford's ^{131}I could be expected to detect any increases in thyroid disease related specifically to Hanford's ^{131}I.

V.17 Did you include all persons that could be found and who volunteered to be included? The statements on page 10 are too vague. Why were 909 potential participants not included? Explain what effect not including those 909 persons might have had on the results.

Response: Every eligible potential participant who was selected for the study, was located, agreed to participate, and attended an HTDS clinic was included in the analysis, except for seven who were nonevaluable according to the study's predefined criteria (see sections IV.B and IX.A of the Final Report). Detailed information is provided in sections V.B and V.C of the Final Report to describe how many of the 5199 individuals originally selected for the study actually participated and the reasons for nonparticipation. A more detailed discussion of the possible effects of nonparticipation is provided in section X.C.1 of the Final Report.

V.18 Lincoln County, with the highest Multiple Sclerosis incidence on the planet, should've been look at too.

Response: As noted above, the counties from which participants were selected were chosen based on the likelihood that residents would have been exposed to atmospheric releases of ^{131}I from Hanford. From the information available at the time it did not appear that Lincoln County residents would have likely received as much exposure as residents of the counties that were selected. As discussed above, it is not necessary to include residents of other counties, such as Lincoln, in the actual study in order for the HTDS results to be meaningful to those residents. Also, as described in section II.A of the Final Report, the HTDS was specifically mandated to investigate thyroid disease, not other disorders.

V.19 The report of the HTDS does not include a section on dietary intake methodology, including the strengths and weaknesses of different approaches to obtaining dietary intake information, such as use of proxy respondents and accuracy of reporting intake from many years ago.

V.20 The study methods did not seem to include any attempt to ask questions that would allow for assessment of internal reliability of dietary recall.

V.21 It is difficult to get an overview of who provided dietary information for what proportion of the final HTDS study population. We urge you to summarize this information in Section V or Section VIII.

V.22 While you assessed thyroid disease using exposure based on both the reported dietary intake and the HEDR reference diet, you did not seem to include any comparisons of the reported diet to the reference diet.

Response: Obtaining information about specific aspects of each participant's dietary intake when he or she was an infant and small child, more than forty years after the fact, was a major challenge faced by the study. An extensive discussion of alternative methods was not included in the report because there were few feasible options. A discussion is presented of the need to interview proxy respondents, because the participant him/herself would have been too young at the most relevant times to remember (see section V.D.1.a of the Final Report). A rather extensive discussion is also presented to describe the special attempts made to modify the more standard approach of interviewing typically used in epidemiologic studies to include elements based on principles of cognitive interviewing to enhance memory and recall (see section V.D.2.a of the Final Report). Assessment of the internal reliability of the dietary questions, using such standard techniques as re-interviewing, was not deemed appropriate under the unique circumstances of this study and the data collection methods used, and would not have provided very informative results.

The proportions of living evaluable study participants whose doses were estimated from dietary and other data provided by their CATI respondents are summarized by in-area status in Table IX.B-1 of the Final Report. The relationships of CATI respondents to their corresponding participants are summarized in Table V.D-5 of the Final Report.

Comparisons between dietary data reported by CATI respondents and HEDR defaults were not performed because the two were derived from widely different sources. As described in section V.D of the Final Report, reported dietary intakes were obtained from CATI interviews of respondents with direct personal knowledge of their corresponding participant's life during 1944-57. The CATI respondents were mostly elderly, most commonly the participant's mother, and were asked to recall information from a period 35-50 years before the interview. In particular they were asked to provide point estimates of the quantities of food and milk products consumed during that period by the study participant, and by the participant's mother if she was pregnant with or breastfeeding the participant. In contrast, the HEDR default dietary data was derived from data collected during the Nationwide Food Consumption Survey of 1977-78, which were adjusted to reflect food consumption in the period of interest (1945-57), and consisted of empirical distributions of quantities consumed, rather than point estimates. (6).

V.23 *I would consider reviewing all cases of malignancy and a subset of the remaining cases by two or more experienced thyroid cytopathologists. Also, I would ask them to review the cases with hypocellular samples with abundant colloid to determine whether they agree with their classification.*

Response: A comprehensive review of all 259 biopsy specimens (rather than a subset) was done and the results summarized in the Response to NAS document (see Appendix 24). Cases of malignancy could not be reviewed since the HTDS no longer had access to those specimens, however all HTDS diagnoses of cancer were based on separate reviews by the HTDS pathologist of the original pathologist's interpretation. Of 19 cases, there was complete agreement with the diagnosis of thyroid cancer in 17 cases. For one case, the slides were not available but the HTDS pathologist reviewed the initial pathology report, performed by a nationally recognized AFIP pathologist, and expressed confidence in the reported diagnosis. For the final case, there was disagreement between the original pathologist, who did not find thyroid cancer, and the HTDS pathologist, who found a 4 mm focus of papillary carcinoma.

With regard to hypocellular specimens with abundant colloid, all of these cases were reviewed. There was no disagreement by the reviewing cytopathologists in any of these cases to suggest that a neoplasm or carcinoma had been missed.

These results, in addition to the comprehensive review of all FNA specimens (see Appendix 24, Response to NAS), strengthens the validity of the approach in the original study design of having a single, experienced cytopathologist review all of the FNA specimens.

V.24 Page 116. In this section it is stated that "...four participants were determined to be non-evaluable". On p.188 the number of non-evaluables is given as six. I don't know which number is correct, but I believe that the same number should be given in both of these places.

Response: Section IX.A of the Final Report contains information on the number of participants who were eligible but nonevaluable. The correct number is seven. Six did not have complete residence histories for the period from the beginning of their possible exposure to $^{131}I^{-}$ from Hanford through the end of 1957, and the seventh had a tracheotomy tube in place which prevented palpation of her thyroid at her HTDS clinical examination.

V.25 There is no sense of how many nodules greater than 1.5 cm were biopsied or unbiopsied (i.e. ultrasound detected, nonpalpable nodules that remained unbiopsied).

Response: Among participants who had nonpalpable ultrasound-detected abnormalities, 34 had dimensions greater than 1.5 cm in three dimensions. Of these 34 participants, 25 underwent FNA biopsy while the remaining 9 did not undergo biopsy. Those not undergoing biopsy either declined FNA or in a few cases it was not recommended since the recommendation to biopsy such nodules was not instituted until after the first year of the study. For the 25 participants who underwent biopsy, 23 had adequate biopsy specimens. All of these were benign except one, which was suspicious for a follicular neoplasm. There were no cases of thyroid cancer.

V.26 Data can also be presented starting with the nodules that were biopsied and determine how they were discovered (palpation, ultrasound) and show the pathological results and final assessment.

V.27 I believe that it would be very helpful to have tables that summarize "paths to diagnosis". I believe that there should be about three of these tables, one for cancer, one for benign, and for all nodules.

Response: An analysis of the “pathway to diagnosis” has been performed for all of the diagnoses made by FNA biopsy in the HTDS (see sections IX.C.1.a and IX.D.1.b). Tables have been provided for thyroid cancer (Table IX.C-3) and for benign nodules and nodules suspicious for follicular neoplasm (Table IX.D-9) These analyses, done on 256 participants who had FNA procedures performed, are discussed in these sections of the Final Report as well as in the Discussion (X.C.2).

V.28 The FNA results are somewhat unclear in the draft report. For instance, on page 8 of the executive summary it states that 259 participants had FNA. 62 (24%) were recommended to have further biopsy but in the post-clinic medical records review it states medical records documenting further diagnostic studies were requested for 35 participants (page 9 of executive summary). What happened to the other 27 participants (62 less 35)? Are these the patients who have a nodule which is suspicious for malignancy or neoplasm (Background, page 15)? But this number appears to be 16 (the difference between line 2 on Tables VIII-18 and VIII-22 and -23).

V.29 Sixty-two of 259 FNA (24%) were recommended to have further biopsy or surgery. Twelve of the participants had thyroid cancer and 14 follicular adenoma. This leaves 13% of the FNAs needing further evaluation. The report should go through these cases and show what percent of cases actually needed confirmation due to the uncertainty of the FNA/pathologist. They should compare the nondiagnostic percentage in this study to published thyroid FNA results.

Response: These comments concern two related issues: the completeness of follow-up for persons having FNA and the percentage of nondiagnostic FNA procedures. Regarding the first issue, of the 259 participants who underwent FNA, 47 (not 62) were recommended to have further biopsy or surgery. (The 62 referred to in the Draft Report also included individuals who were to be followed for nuclear medicine scans, additional lab work, etc, but not exclusively because of FNA results; see section V.H.3.a of the Final Report). Of the 47 who were recommended to have further biopsy or surgery, 30 did so and were consequently diagnosed with thyroid cancer (12 cases), follicular adenoma (5 cases), or benign nodule other than follicular adenoma (13 cases). The remaining 17 participants were classified as Suspicious for Follicular Neoplasm because they did not go on to have further biopsy or surgery. See sections VIII.B.3.c and IX.D.1.a.1 of the Final Report for further discussion of this category.

The second part of this question refers to nondiagnostic FNA rates. In interpreting these data, it is important to note that the amount of specimen material obtained by the HTDS physician team exceeded what is typically obtained in clinical practice. For example, the number of aspirations per nodule typically was between 4-10 with 10-20 slides of material provided to the cytopathologist for review. The number of nondiagnostic aspirates (hypocellullar or acellular specimens) was 7 (2.7% of the 259). In addition there were 18 persons with a single nodule for which the FNA showed abundant colloid but hypocellularity. These were classified as benign. This classification was made in part due to the confidence of having extensive sampling as noted above in this study. However, it is acknowledged that a suspicious or malignant lesion cannot be entirely ruled out in such cases. Thus, if we assume these 18 cases to also be nondiagnostic and add them to the 7 true nondiagnostic cases noted above, this yields 25 cases (9.7%) as the upper bound of inadequate or nondiagnostic specimens in this study. Thus, the rate of nondiagnostic FNA specimens in this study is 2.7-9.7%, a figure that is well within the 2-20% figure in the published literature.

V.30. The Discussion talks about thyroid neoplasia (Discussion, page 15) and the Utah study, but are the results comparable? Is it reasonable to perform a dose response analysis on different combinations of neoplasms? That is 1) adenomas and carcinomas together, excluding colloid and non-neoplastic nodules and 2) adenomas and carcinomas and nodules suspicious for malignancy, excluding colloid and non-neoplastic nodules?

Response: We understand this to be a question about whether the HTDS performed an analysis identical to the one performed in the Utah Study which showed a statistically significant dose-response for thyroid neoplasia (thyroid cancer plus benign follicular neoplasms). The HTDS undertook an even more comprehensive approach to define alternative definitions of thyroid outcomes in order to determine if a true dose-response relationship might have been missed in the primary analyses. These alternative outcome definitions included the following: benign nodules plus nodules suspicious for follicular neoplasm, benign nodules excluding non-neoplastic disease (such as Hashimoto's or Graves), benign nodules detected only by palpation (prior to ultrasound review), benign colloid nodules (see section IX.D.1.a of the Final Report). In addition the HTDS investigated thyroid neoplasia, which was defined as all thyroid cancer and all benign follicular neoplasms (excluding colloid nodules). This was identical to the outcome of thyroid neoplasia defined in the Utah study. Dose-response analyses were performed on all of the above alternative outcome classifications (see section IX.E of the Final Report).

V.31 To confirm the findings of a lack of a dose response relationship, and to investigate further the role of palpation and ultrasound, I suggest carrying out separate analyses for the endpoints of palpable nodules and ultrasound nodules larger than 1.5 cm in average dimension.

Response: We interpret this comment as requesting one dose-response analysis for palpable nodules and a second dose-response analysis for ultrasound nodules (palpable and nonpalpable) which are greater than 1.5 cm in average dimension . The dose-response analysis for "Any solitary palpable nodule" can be found in Section IX.F.1.a.1 and IX.F.2.c.1). Regarding the ultrasound nodules, a comprehensive analysis has been performed for focal ultrasound-detected abnormalities by size: one for those greater than 5mm, one for those greater than 10 mm, and one for those greater or equal to 1.5 cm in average dimension. These analyses can be found in Section IX.P.1.a.1 and IX.P.2.b. In addition, a separate analysis (see Section IX.D.1.a.3. and IX.D.2.c.3) has been done for benign nodules detected by palpation (excluding those found with the assistance of ultrasound).

V.32 I found that one table in the analysis plan that was not in the final report. This was the number of palpable nodules that were not confirmed by ultrasound. Since the analysis plan was reviewed and agreed to by several groups, it is best to complete all aspects.

Response: This table has been added to the Final Report (section IX.D.1.a.3)

V.33 I believe that the description of multinodular goiters and multinodular glands and their overlap with the category of thyroid nodules should be made clearer.

Response: As described in section IV.C of the Final Report, the only difference between multinodular gland and multinodular goiter in that the estimated size of the gland in the latter category is greater than 2-fold enlarged. Dominant palpable nodules in a multinodular gland or goiter were biopsied in the same manner as solitary thyroid nodules.

V.34 The laboratory findings should be scanned to be certain that the appropriate units are stated whenever a value is given. Also, while the various TSH methods are described, it would be helpful to describe them by their level of sensitivity.

Response: The appropriate units have been included in the Final Report. Detailed characteristics (such as sensitivity) of the various laboratory tests used are not provided as they do not alter or influence the criteria for diagnosis for a given outcome.

V.35 Has the incidence of juvenile onset thyroid disease been looked at?

Response: Any participant with a history of thyroid disease, whether as a juvenile or adult, was asked to release medical records. The information obtained was then incorporated into the HTDS results regardless of the age at which the diagnosis was made. Thus, even though the actual age at onset may not be available, the diagnosis of any thyroid disease in a participant is included in the HTDS results.

VI. COMPARISONS

VI.A. Comparisons to literature or other studies related to dose-response, radiation and thyroid disease.

VI.A.1 How do these findings compare with the literature (regarding radiation and thyroid disease)?

VI.A.2. What does the literature say about dose-response?

VI.A.3 Do these results agree with "current knowledge" (about radiation and thyroid disease?)

VI.A.4 How does [the study] fit into the whole spectrum of other exposures of I-131 and thyroid disease?

Response: The answers to these four questions come from hundreds of scientific studies done over the last 75 years on the effects of radiation and thyroid disease. It is well known that some types of radiation exposure can cause thyroid disease. However, the frequency and type of thyroid disease, and age at onset of the disease, all depend on many factors such as the type of radiation exposure, the route and duration of exposure, the age of exposure, and the magnitude of the dose. These and other factors determine the risk, or chance, of actually getting thyroid disease. For some circumstances of radiation exposure, the chance of getting thyroid disease during a person's lifetime may be high while in other circumstances of radiation exposure the risk may be so low that an individual may never experience thyroid disease from that exposure.

Since there is an extensive scientific literature on this subject, the response to these questions will necessarily be abbreviated. In general, the most common types of thyroid disease that have resulted from radiation exposure are benign and malignant thyroid masses. While hypothyroidism can result from environmental exposures, the doses must be exceedingly high to cause this problem. There are relatively few reports that suggest that hyperthyroidism results from radiation exposure. Some studies have suggested that autoimmune thyroiditis, which is quite common in

the general population, may be increased after radiation exposure. However, clear evidence for this is lacking and much further study is needed before this can be stated with certainty.

The literature shows fairly clearly that radiation exposure from external gamma radiation produces a linear dose response. Again this is subject to the many different risk factors that determine whether radiation exposure in a given circumstance actually causes a thyroid disease. For example, it was clearly shown that young people (less than 15 years at the time of exposure) who were exposed to A-bomb radiation at Hiroshima or Nagasaki developed excess thyroid cancer with a linear dose-response. However, those persons who were older than 20 years had almost no risk of thyroid cancer during their lifetime even though they were exposed to the same levels of radiation. The type of dose-response for exposures involving radioactive iodine (e.g., Chernobyl) is much less clear. Additional information on this issue has been added to this Final Report (see Section X.D).

The response to the question of whether the HTDS results agree with "current knowledge" is complex. The HTDS results are "consistent" with the world literature regarding radiation exposure and thyroid disease in the sense that many of the factors that would predict low risk of thyroid disease are characterized by the HTDS cohort of participants. Although the cohort was composed exclusively of people who were in utero, infants, or young children at the time of greatest exposure, and therefore the most likely persons to develop radiogenic thyroid disease, other factors of the exposure might be predictive of low risk. The most important of these factors is probably the low magnitude of the radiation dose. A second factor is that these low doses were accumulated over months or years whereas in other populations exposed to environmental radiation, where the risk of thyroid disease was significant, much or all of the exposure happened acutely, usually over hours. Another factor may be that the exposure from Hanford was almost exclusively ^{131}I and did not include significant external radiation or other types of radioiodine which were present in other population exposures.

VI.A.5 An explanation of the reasons for differences in the findings of NTS and the HTDS would be helpful in assessing the HTDS results.

Response: With regard to the question of NTS exposures we interpret this question to mean the Utah Study, which evaluated the risk of thyroid disease from exposures from the Nevada Test Site. We have provided additional discussion in the Final Report (see section X.D) regarding the differences in both the type of exposures in these two studies, as well as the differences in results.

VI.A.6 Questions about the correlation to Dr. Rudy Nussbaum's 801 health surveys for downwinders.

Response: Results of scientific investigations depend greatly on study design. Different study designs may produce different results, even though each study is trying to answer the same scientific question. The "801 Health Survey" was based on completed questionnaires received from approximately 800 individuals who responded to a general request for people who considered themselves to be Hanford Downwinders to fill out a questionnaire. Surveys of this type can be very misleading (i.e., biased) if those who choose to reply are systematically different with respect to their disease status and their exposure experience. For example, if those who have thyroid disease and who were exposed to Hanford radiation are more likely to participate than those who don't have thyroid disease (but were also exposed to Hanford radiation), the results will incorrectly indicate that there is an association (a relationship) between exposure and thyroid disease. The HTDS was designed to limit this kind of bias by following an entire cohort of people

over a long period of time to determine the likelihood that those exposed to Hanford radiation were at an increased result of developing thyroid disease as a result of their exposure. A survey like the "801 Health Survey" is not able to answer this question.

VI.A.7 This study seems to assert that its results are superior to other work that has been done on I-131.

We do not know the origin of the comment this question is referencing. With regard to studies of Hanford ^{131}I exposures, the HTDS is the only study of its kind that has ever been conducted. With regard to any study of ^{131}I, the HTDS is unique in terms of the degree of peer review and quality control that has characterized the study, and has been conducted in as rigorous a manner as any other published study.

VI.A.8 What is the incidence of thyroid cancer in all persons born in the US between 1940 and 1946?

There are no data on the incidence of thyroid cancer in the United States as a whole. The only source of population-based cancer incidence data is the Surveillance, Epidemiology and End Results Program of the National Cancer Institute. This program consists of approximately ten cancer registries in locations throughout the United States. Although the areas covered include urban and rural areas in all different regions of the U.S., they are not necessarily representative of the overall U.S. population. The program did not begin operation until 1973 in some areas and 1974 in other areas. Thus, it is not possible to determine the exact answer to this question. In general terms, the incidence of thyroid cancer in adults in the U.S. (persons born between 1940-1946 would be adults now), based on recent SEER data, is approximately 4 per 100,000 per year. The incidence is higher in females than males (approximately 6/100,000 per year vs. 2/100,000 per year, respectively) and increases with age. Thyroid cancer under age 20 is very uncommon in the U.S.

VI.A.9 Did any of the 19 persons in the study identified as having thyroid cancer know they had thyroid cancer in advance of the study?

Response: Of the 19 persons identified with thyroid cancer in the HTDS, 7 had the diagnosis and treatment prior to entering the HTDS, 12 persons had the diagnosis made by HTDS physicians. In addition, one individual reported a prior history of thyroid cancer for which medical records were unavailable. Therefore, 8 of 20 persons with a diagnosis of thyroid cancer knew of the diagnosis prior to their participation in the HTDS (see section IX.C.1 of the Final Report).

VI.A.10 Question about whether 2,000 people with thyroid disease is exceptional, normal or below normal.

VI.A.11 Is there a higher percentage of thyroid malfunction in Hanford than nationally?

VI.A.12 The population also had a surprising amount of thyroid disease, although its prevalence was not dose related. These findings raise questions that should be answered with regard to the population studied, although it is hard to see how significant bias could be introduced considering the way in which the population was selected.

Response: With regard to the above three questions referring to comparison of thyroid disease in the HTDS cohort with that nationally, an entire section has been added to the Final Report which provides a review of the world literature on the prevalence of thyroid diseases (see section X.E). This section compares the results from that literature to the results obtained from the HTDS.

VI.A.13 It is surprising that the presumably normal study population had such a large number of prior diagnostic X-ray exposures. That 36% of the population had prior upper GI series and 34% had X-rays of the head, to say nothing of the 24% of the population that had a CT scan of the upper body raises a question as to whether the selected study population is truly representative. The availability of an appropriate control population for comparison would add another layer of reassurance. And informal and anecdotal query of a limited number of diagnostic radiologist colleagues suggests that these numbers are beyond what might be expected in ordinary practice with a "normal" population.

Response: The HTDS did not evaluate the actual number of x-ray procedures but rather whether or not a participant or his/her CATI respondent reported that the participant had any history of such exposures. Independent validation of these self-reports was beyond the scope of the HTDS. Furthermore, there are no studies of the frequency of use of x-rays and other procedures in unselected populations from that time period which might provide a suitable comparison. However, the question of whether the HTDS population is "representative" regarding the use of such procedures is not of direct relevance in evaluating the dose response with Hanford radiation. The more important question is whether the reported frequency of the procedures among study participants confounded or modified the effect of Hanford radiation. This was evaluated formally in the analyses of each outcome, and there was no evidence of confounding or effect modification according to the number of self reported medical radiation exposures (see the sections entitled "Confounding and Effect Modification" in sections IX.D through IX.M, IX.O, and IX.P of the Final Report).

VII. MORTALITY STUDY AND ANALYSIS

VII.1 I trust your study is also covering those people who lived in the area but who have died.

VII.2 These [deceased] lost souls most definitely need to be included in all these studies.

Response: As described in section V.B and Appendix 23 of the Final Report, the HTDS attempted to identify all potential participants who were deceased among the original 5199 identified for study, regardless of where they resided when they died. The study also requested a

copy of the death certificate for all those confirmed as deceased. Of the 543 individuals confirmed as deceased, a death certificate was obtained for 504 (93%).

VII.3 *I noticed that 541 people were deceased. Please tell me how that compares with traditional mortality tables, and should I be concerned? As I looked at the tables, it seems that the risk of thyroid problems was actually reduced for those who were exposed.*

VII.3 *It is of interest and a little disturbing that mortality in the cohort was 20% higher than expected. The primary contributor to this number was apparently congenital abnormalities and perinatal problems, which amounted to an excess of over 2 times that, expected. This could hardly be a radiation effect since many of these births occurred before Hanford went into operation, but further discussion would be helpful.*

VII.4 *They have one hit - higher mortality. How does the study deal with this? They shush things. They say there were more dead people before and afterwards so they won't really worry about it.*

VII.5 *Page 6 of the Summary Final Report: Researchers didn't do any studies to make assertion that "although there is a high death rate and although the reasons for this higher death rate are not known, it is not likely that it is related to radiation from Hanford."*

VII.6 *The impression being given is that radiation is not a concern for Downwinders. I am concerned that they do not know why people died, they might have had thyroid cancer.*

Response: Based on the information obtained regarding cause of death from the death certificates for deceased potential participants, an analysis was conducted to investigate whether the mortality experience in the HTDS cohort overall was unusually high, relative to what would be expected based on the mortality experience of the population of the same region over the same time period. Additional analyses were conducted to determine whether there was any indication of an excess in mortality in the HTDS cohort from conditions that might be related to one or more of the primary outcomes of interest regarding thyroid disease. In summary, there was no overall increase in total mortality over what would be expected based on the mortality experience of the population of Washington State during the same time period. This was true for both men and women. However, there was an excess in deaths due to conditions of the perinatal period, which was found in both men and women.

Findings based on preliminary analyses which were included in the Draft Final Report (January 1999) indicated a similar excess in mortality due to conditions of the perinatal period, and also suggested a 20% excess in mortality overall. However, those findings were based on a more crude analysis which included in the cohort only those who attended a HTDS clinic, as well as those who died. The present analysis is more complete, and includes all those located from the original cohort of 5199 individuals, regardless of whether they participated in the study or not (and including those who died).

A detailed description of the methods used to assess mortality in the HTDS cohort and the full results of these analyses are presented in Appendix 23 of the Final Report. In addition, the Discussion section (Section X.C.1) has been expanded considerably to consider the possible impact on the radiation dose-response of deaths in the cohort.

VII.7 Year of birth as an indicator of exposure can be somewhat misleading when is dealing with congenital anomalies or pregnancy complications. Exposure at conception or in early pregnancy may be relevant, and this might be in the previous calendar. If you have the actual birth dates, you could estimate year of conception and examine the data that way. This would be more appropriate analysis.

Response: Year of birth was used in an attempt to see whether the observed excesses in mortality were concentrated among persons born around the time of the peak releases from Hanford (i.e., 1945 and 1946). A number of analyses were repeated separately for the birth cohorts defined by the period 1940-44, and 1945-46 to reflect what might be reasonably assumed to be different exposure conditions. However, it is well recognized that this is a very crude approach to assessing exposure to Hanford radiation. It was not possible to conduct dose-response analyses based on individual estimates of exposure for persons who had died. Thus, the mortality analyses conducted using cause of death information were not capable of formally assessing the relationship between Hanford radiation exposure and outcomes such as congenital anomalies or pregnancy complications. The HTDS was not designed to evaluate mortality, and these analyses were never intended to investigate a relationship between Hanford radiation dose and cause of death among those in the cohort who died.

VII.8 Two findings of mortality analysis are especially notable: the excess in perinatal mortality and in fatal congenital anomalies, and the particularly high mortality among those born in Franklin County. To what extent do these overlap? i.e., Is the excess in Franklin County due to perinatal mortality? This should be addressed.

Response: As noted above in response to comment VII.7, the HTDS was not designed to evaluate mortality in this cohort. The study was not conducted in a manner that would allow for a detailed analysis of cause of death, or that would be capable of determining whether mortality in this cohort was associated with radiation exposure from Hanford. The primary purpose of reviewing death certificates for those who died was to determine whether any of the deaths were due to thyroid disease. The primary purpose of comparing the mortality experience in the HTDS cohort to that of the population of Washington State during the same time period was to see whether mortality in the HTDS cohort was substantially different (either higher or lower) than what might be expected based on the surrounding population. Extending the analyses to explore detailed patterns in specific subgroups of the population (e.g., those born in Franklin County and deaths due to specific causes) are not appropriate and are beyond the uses for which these data were intended.

VII.9 Mortality results should be discussed in the context of other relevant studies such as Sever et al. (Am J Epidemiol 1988; 127). Even though HTDS wasn't designed to explore mortality findings in detail, they deserve more consideration than is currently given on p.8 of Section IX.

Response: As noted in the responses above, the results of the mortality analyses conducted as part of the HTDS were not intended to address the same types of questions that studies like those of Sever et al were, and are not capable of doing so. Thus, direct comparisons of the HTDS mortality results are not appropriate. Nevertheless, in response to a number of comments and suggestions received after the release of the Draft Final Report, the mortality findings are

presented and discussed in considerably more detail in Appendix 23and section X..C.1 of the Final Report.

VII.10 The particular perinatal conditions and congenital defects should be described. This could prove informative, especially if coupled with data on place of birth and year of conception.

Response: This detailed information was not always available on the death certificate, and there were too few deaths from any specific cause to allow for a meaningful analysis. That is why such deaths were grouped into categories of similar causes. Further, as indicated above, this study was not designed to formally assess the relationship between Hanford radiation exposure and specific causes of death, nor was it capable of doing so.

VII.11 In terms of the excess in cardiovascular mortality, are there any obvious differences in risk factor prevalence that would explain it?

Response: Such an assessment is not possible, given the data available and thus is beyond the scope and capability of the HTDS.

VII.12 The SMRs were calculated using Washington State as the standard. Are there any regional data available that would provide a better comparison or are the statewide data correct?

Response: In conducting an analysis of this type, the most important consideration is whether the comparison population (and mortality rates used to calculate expected numbers of deaths) are truly comparable to the population under study (in this case the HTDS cohort). In addition, a practical limitation is often encountered in terms of the availability of the detailed data needed to perform the calculations. For the present analysis, it was necessary to have access to mortality rates by age, sex, race and geographic area over a period of approximately forty years or more by at least major category of cause of death. This dictated that we used statewide data. Such detail was not available from a smaller geographic region. In our judgment, the data for the State of Washington constituted the most comparable data available. Potential limitations of this approach are addressed further in the Discussion section.

VIII. COMMENTS REGARDING FINDINGS, INTERPRETATIONS, AND CONCLUSIONS

VIII.1 What is the "bottom line" of the study? What does it show?

VIII.2 The excerpt on page 7 [of Summary Report] states that 'the Director shall conduct a study of thyroid morbidity of the population.' A morbidity study determines the number of cases of a particular disease occurring in a given number of population. You chose to go further and add the objective of determining whether disease was increased. You should either not make that conclusion or should also address the other two possible conclusions of a morbidity study by also forming conclusions on whether there was no effect or that there was a decrease in disease as the radiation dose increases.

VIII.3 How can you explain away a problem that effects so many people - multiple family members with thyroid disease (no family history)?

VIII.4 Your findings basically say, hey if you "lived off the land" as our family did in North Idaho, you received no radiation-nothing to be concerned about. If you happen to have thyroid problems or any other problems health wise there is no way it could be tied to the Hanford releases.

VIII.5 What about long-term effects?

VIII.6 The Hanford Environmental Dose Reconstruction Project did not release their findings until April 21, 1994 and only then for representative doses, not individual doses. It is my opinion that chronic long-term exposure to Iodine 131 in the air, in the water, in the soil, in the food, in the milk, in whatever dose, resulted in thyroid disease.

VIII.7 I question why the study looked for a dose-related effect? The information I have from this population is that there is disease, with a wide variation in exposure and dose. Science may appreciate knowing how dose-response to disease was found by the Hanford Thyroid Disease Study. The Downwinders I have spoken with, know that the study does not reflect the disease they have experienced.

VIII.8 It is not possible to say that the thyroid disease in this population is not related to the Hanford emissions. There are health effects in this population that the design of the study does not address. The Downwinders are not reassured that the emissions from Hanford did not contribute to their thyroid disease. With all due respect to the researchers, the results of the Hanford Thyroid Disease Study are not conclusive, and do not accurately reflect the numbers of persons with thyroid disease and other diseases in the Hanford population.

VIII.9 .There should be discussion of: interpretation of the meaning of negative findings in an epidemiologic study

VIII.10 My wife, who was born in 1944 and was a downwinder, died in 1993. Of her graduating class (100 students) she was the 8th one to die of cancer…that is just not normal. Others in her family also have weak thyroids. I am troubled by the statement that there is no link, something must be wrong.

VIII.11 Are the researchers saying that 700,000 - 800,000 rads of I-131 is not harmful to the public?

Response: The HTDS was designed to determine whether exposure to atmospheric releases of primarily ^{131}I from the Hanford Nuclear Site between 1944 and 1957 resulted in increased thyroid disease among those exposed. The primary objective of the research was to describe in what way any increase in thyroid disease observed is related to the dose of radiation received; that is, to describe the characteristics of any dose-response relationship. This is the best way to assess whether exposure may have caused disease. The study was conducted as a long-term follow up study over a period of more than forty years after exposure. That was done to capture as much as possible any long-term or late effects of radiation exposure. The primary analysis utilized an

estimate of thyroid radiation dose for each individual based on information about their residence history and dietary consumption patterns during the times of the Hanford releases.

The overall ("bottom line") result is that this study found no evidence in any of the analyses that increasing dose to the thyroid from Hanford radiation was associated with an increased cumulative incidence of any of the disease outcomes or with increased prevalence of thyroid ultrasound-detected abnormalities, with results of thyroid laboratory tests, or hyperparathyroidism. These results remained the same when alternative methods of assessing radiation dose were used, and after accounting for uncertainty in dose estimation. There is no evidence that the absence of a dose-relationship was due to bias in selection of the cohort, loss to follow-up, or enrollment and participation.

Very important in the interpretation of these results is the assessment of the ability of the study to detect an increase in disease risk if it is present (i.e., the statistical power of the study). In order for the findings of a study showing an absence of an effect like that seen in this study (e.g., a negative study) to be very meaningful, there must be adequate statistical power to detect an effect of the magnitude that might be expected based on existing knowledge. The projections of study power, which were based on the results of the Pilot Study, were actually exceeded in the Full Study (as shown in Table IX.B-14 in section IX.B). Nevertheless, because uncertainties in the individual dose estimates could be expected to reduce study power, we undertook additional analyses to estimate the impact on study power of incorporating such uncertainties in the dose estimates. These new analyses are described in section IX.B.4. Although the effect of dose uncertainty was, as expected, to reduce the statistical power of the study, the reduction was modest. Even after accounting for uncertainty in doses, the HTDS had greater than 80% power to evaluate each of the hypotheses originally specified.

Given the principal differences between the radiation exposure circumstances at Hanford and those of other populations studied in relation to radiation-induced thyroid disease, the findings of this study are not inconsistent with the current published literature regarding the effect of exposure to radioactive iodine and the risk of thyroid and parathyroid disease. This is particularly so given the relatively small magnitude of the estimated thyroid radiation doses in members of the HTDS cohort (mean = 174 mGy) and the relatively protracted nature of the exposure over time. There is little evidence in the literature to suggest that persons exposed to radioactive iodine at the levels found in this study over a period of months or years would experience higher rates of thyroid or parathyroid disease as a result of their exposure.

This is not to say that there isn't thyroid disease in the population exposed to the Hanford radiation or in the HTDS cohort, or that exposure to radiation isn't harmful. The HTDS results show that thyroid disease is present in this cohort, and the results of the dose reconstruction project show that cohort members were exposed to Hanford radiation. It simply says that there is no evidence in this study of a link between exposure to Hanford radiation and the subsequent development of thyroid disease. This has raised a question for many of whether the study was incapable of finding that link because of the uncertain nature of the dose estimation used in the primary analyses and a concern that such uncertainty is so great that it renders the quantitative dose-response results inconclusive. The study has attempted to address this possibility in three ways. First, alternative qualitative methods of assigning exposure were used. Results from these analyses were consistent with those from the quantitative dose-response analyses. Second, two different approaches were employed to evaluate the impact of dose uncertainty on the primary risk estimates. Neither resulted in findings that were materially different from those ignoring such uncertainty. Third, the impact of dose uncertainty on study power was assessed using simulation methods. These analyses revealed that any reduction in statistical power due to uncertainty in

dose estimation was modest, and that even after accounting for such uncertainty the study had adequate statistical power to detect effects as small or smaller than those in the existing published literature.

Although any epidemiologic study is limited to some extent by uncertainty in the assessment of exposure, the impact of such uncertainty on the power of the study and the estimation of risk is seldom addressed to the extent attempted here. Further, the fact that epidemiologic investigations are inherently "uncertain" does not imply complete randomness or unpredictability, nor does it mean that reasonable conclusions cannot be drawn from such studies. Although these findings do not definitively rule out the possibility that Hanford radiation exposures are associated with an increase in one or more of the outcomes under investigation, the power of the study, even after accounting for the uncertainty of dose estimates, suggests that a failure to detect such an effect, even if it is very small, is unlikely.

VIII.12 I urge you to consider neck x-rays for people who were conceived in the Hanford area. There are many problems associated with the Klippel-Feil Syndrome–many are hidden symptoms because we do not communicate the symptoms. Since we are born with this problem many of the symptoms are normal to us, so we do not communicate them to a Doctor, until it is too late. Then we have paralysis, nerve problems, stenosis.

VIII.13 There are no comprehensive studies on the very specific area where it truly would have affected people– namely the unborn babies, and newborns of the workers at Hanford. (The workers took home a higher concentration of radioactive particles). They should be looking for spine deformities, mental retardation, and growth pattern problems specifically of the children of the workers of Hanford, or those conceived in the Hanford area.

VIII.14 M.S. is the highest in the nation in the beautiful Northwest, which is where Hanford Nuclear Reservation is located.

VIII.15 Health problems to Downwinders include chemicals, nuclear reactors and pesticides. Several relatives and neighbors are sick with various diseases including leukemia, MS and thyroid cancer.

VIII.16 Are other studies going to be done for other diseases?

Response: As described in section II.A of the Final Report, the HTDS was mandated and funded by Congress to specifically investigate whether thyroid disease was increased as a result of Hanford radiation releases from 1944-1957. Although we understand that there is considerable interest in studying the possible effects of Hanford exposures on diseases and conditions other than thyroid disease, such as multiple sclerosis, it was beyond the scope of the HTDS to do so.

VIII.17 I have followed the study with great interest particularly after I was diagnosed with papillary thyroid carcinoma in April 1998. My thyroid was located behind the sternum. Therefore the carcinoma had not shown up as a nodule in the neck during routine physicals.

VIII.18 Did the HTDS collect information about the location of the thyroid? Did such information show a greater incidence of carcinoma and other thyroid diseases? If so, I would like to see this information highlighted in the study or in a separate finding. Are there other studies that address the location of the thyroid?

Response: The HTDS did not collect information that would allow the investigators to distinguish anatomical differences in the location of the thyroid gland. Rarely, thyroid enlargement can occur in unusual locations (for example, behind the sternum). Most of these conditions are benign, although even more rarely it is possible for a thyroid cancer to occur there. However, in such instances a person would likely develop symptoms that would lead to medical care. Since we sought to obtain prior medical records for all participants reporting prior thyroid medical problems, it would be highly unlikely for such a condition to have been missed by the HTDS evaluation process.

IX. GENERAL COMMENTS AND COMMUNICATIONS

IX.1 I think my grandchildren are victims of Hanford. They have all kinds of cancer in their lives.

IX.2. Science is useless in solving social problems.

IX.3 The federal government conspired to eliminate liability for the releases.

IX.4 This study does not affirm my experience with the thousands of Downwinders with whom I have spoken who call the Network to talk about their thyroid and other diseases.

IX.5 No study has been done with a population exposed to constant radiation in varying amounts over a long period of time. Neither has there been a study that can account for each individual response to a stimulus.

IX.6 There is no consideration for political context in the study. The DOE and the US Government are political entities, and they could not do a scientific study - they could get a technological answer, but not a scientific one.

IX.7 .The Green Run of December 2, 1949 was an immoral act and yet the government has never apologized.

IX.8 There was no provision in the study to cover lost wages during the testing. I just couldn't afford to take off from work.

IX.9 The thought of government people chopping pieces off my thyroid sent a cold chill up my spine.

IX.10 Request that the FHCRC publicly retract the statement: "These results provide rather strong evidence that exposures at these levels to I-131 do not increase the risk of thyroid disease or hypoparathyroidism. These results should consequently provide a

substantial degree of reassurance to the population exposed to Hanford radiation that the exposures are not likely to have affected their thyroid or parathyroid health."

IX.11 The HTDS lacks humanity and compassion.

IX.12 The HTDS has effects on emotions and litigation. We (the HTDS) doesn't recognize this.

IX.13 Given the following quote, how could you justify releasing the HTDS draft report when you knew that not all of the final analysis plan had been completed, namely the incorporation of the dose uncertainty into the dose-response analysis? The Study Management Team (SMT) "consider that incorporating the adjustment for dose uncertainty is an essential requirement for the study. That this is indeed a matter of practical importance can be seen from the results of the Utah thyroid study, in which the magnitude of the estimated dose response was roughly tripled by the adjustment for dose uncertainty." The quote appears on page10 of the attachment to the 06/30/97 analysis plan. The attachment is titled "Hanford Thyroid Disease Study Analysis Plan: Summary of Revisions of 1/27/97 Draft, June 30, 1997.

IX.14 Both CDC and the Fred Hutchinson Cancer Research Center (FHCRC) should offer a prominent public apology for their inappropriate characterizations of the power of the study's conclusions during the January 1999 briefings and announcement (see previous comment).

IX.15 The public impact aspects (the way the HTDS results were released) need to be included in their report. Prior to the results on January 28, 1999, I and the media and Congress were completely shut out.

IX.16 The researchers are not dealing with how the report affected the public, that they were ignoring that aspect of their responsibility to handle the social side of releasing the report.

Response: A number of comments such as those above have made it clear that some individuals believe strongly that Hanford radiation emissions have caused their own thyroid or other health problems, or are responsible for a variety of health conditions in friends or relatives. These beliefs are not based on the results or conclusions that have arisen from any scientific or medical studies, but rather are based on personal experience and perception. A number of these comments go on to criticize the HTDS study team for not being sensitive to their health problems or their concerns about the effects of being exposed to radiation from Hanford, and to the way the draft results of the study were communicated to the public.

We understand that it is difficult to accept the results of this study under such circumstances. We respect the rights of all individuals to hold and voice their own opinions and beliefs in this regard, even when those beliefs may not necessarily be based on objective or scientific results. In the same spirit, it is important for members of the public to understand that the primary responsibility of the HTDS team has been to conduct the very best scientific study possible, using the most rigorous scientific methods available. We have been uncompromising in attempting to uphold the very highest standards of excellence in all aspects of the project, and to conduct the study in an unbiased and neutral manner. To help us in this process, we have sought and received extensive

feedback from scientific peers, the federal HTDS Advisory Committee, CDC staff and consultants, and the public in each stage of the study.

Thus, even though the results may be different from what some feel they should be, and no single epidemiologic study ever provides an answer with 100% certainty, we believe we have provided the public with the best possible answer that science could provide to answer this specific question, and have done so in a scientifically rigorous and unbiased manner. Accordingly, we believe it was important to present these findings in a straightforward way, and to provide our best assessment of what they mean and our confidence in them. It is regrettable that this approach was interpreted by some to indicate a disregard on our part for individual circumstances and a lack of compassion and humanity. That was never the intent. We fully realize the potential impact of these findings on individuals, and believe that one of the best ways to show compassion under such circumstances is to deliver the very best scientific product possible.