MALIGNANT MELANOMA AT A SCIENTIFIC LABORATORY

A Synthesis of Reviewer's Comments on the Austin and Reynolds' Study of Employees at the Lawrence Livermore National Laboratory

by

Carl M. Shy, M.D., Dr.P.H.[1]
Harvey Checkoway, Ph.D.[2]
Elizabeth G. Marshall, B.S.[3]

November 15, 1985

From the Occupational Health Studies Program
Department of Epidemiology
School of Public Health
University of North Carolina at Chapel Hill
Chapel Hill, N.C. 27514

[1]Professor, Department of Epidemiology
[2]Research Assistant Professor, Department of Epidemiology
[3]Doctoral Student, Department of Epidemiology

Table of Contents

EXECUTIVE SUMMARY

From the opening of the Lawrence Livermore National Laboratory in 1952 through 1971, the incidence of malignant melanoma among Laboratory employees was unremarkable and equal to expected numbers, based on the frequency of melanoma cases reported in Alameda County. However, between 1972 and 1976 14 new cases were reported in the employee population, and since this number was a significant increase over past experience, the Laboratory commissioned Dr. Donald Austin, Director of the Resource for Cancer Epidemiology Section in the California Department of Health Services, to conduct several epidemiologic studies, the first of which was initiated in 1976-1977 and culminated in Report No. 1 of April 1980. This report indicated that the rate for malignant melanoma at the Laboratory was 3-4 fold above the rate in the surrounding community. An interim report, Report No. 2, followed in February 1983 and showed that the incidence of other cancers was not in excess among Laboratory employees. Lastly an in-depth epidemiologic study of malignant melanoma among Laboratory employees was commissioned and initiated in 1980, and this study emanated in Report No. 3, dated July 3, 1984, by Donald F. Austin and Peggy Reynolds entitled "A Case-Control Study of Malignant Melanoma Among Lawrence Livermore National Laboratory Employees". Seven health scientists, including biostatisticians, clinicians, and epidemiologists were subsequently commissioned by the Laboratory to review the Austin and Reynolds' Report No. 3 and comment on its substantive findings and the methods used to reach its conclusions. Lastly, the present report was commissioned by the Laboratory to synthesize the comments of the seven reviewers and draw conclusions based on these comments.

In approaching this task, we have organized our comments around four recurrent and central issues in the reviewers' comments:

(1) Is there a real excess of malignant melanoma at the Laboratory?

(2) If there is an excess, can this be attributed to occupational exposures at the Laboratory?

(3) Were non occupational factors adequately addressed in this study?

(4) Were appropriate methods and analyses used?

Issue No. 1: Is there a real excess?

While more malignant melanoma cases have been observed in the Laboratory population than in the surrounding community since 1972 - a three- to four-fold excess has been calculated - it is distinctly possible that some or all of the excess cases may be explained by intensive surveillance of suspicious cutaneous moles, a high rate of biopsy of these moles, and an enhanced concern by the Laboratory employees about the reported melanoma excess. This line of reasoning is supported by several observations. Firstly, the increase in melanoma cases between 1972 and 1976 may itself have been a chance temporal cluster of melanoma cases. Whereas 14 new cases were reported, about 5 would have been expected based on incidence rates in the community. Given the higher educational level and superior quality of medical services in the Laboratory population, some excess over community rates would be anticipated. Secondly, following the widespread publicity over the melanoma issue, a spike in new melanoma cases is observable in 1977, and again in 1980 when first reports of the study by Dr. Austin were released. Thirdly a study by Hiatt and Fireman (1984) noted more biopsies for skin lesions among Laboratory employees who were Kaiser Health Plan members than among controls from the same health plan but not employees of the Laboratory. Fourthly, melanomas diagnosed after 1976 among Laboratory employees were distinctly more superficial and more localized (in situ) than the 1972-76 group of melanomas, demonstrating a shift toward earlier detection of lesions in the later years, 1977-1983. If intensive surveillance of cutaneous moles increased the likelihood of finding early stage melanomas, one would expect an eventual harvesting of lesions that would normally be manifested later in time, and a subsequent decline in melanoma incidence. This has not occurred to date. But, if early stage melanomas progress very slowly or even regress spontaneously, then it is possible that many of these early stage lesions would not have been biopsied or clinically detected during the time interval of this study. In this case, the intensive surveillance itself would increase the observed number of melanomas in the absence of a physical or chemical causal agent. Spontaneous regression of moles between juvenile and adult years is a clinically observed phenomenon. Whether early stage melanomas progress slowly or actually regress is unknown and difficult to study, because the lesion is both diagnosed and removed by biopsy. The possibilities that early stage melanomas undergo a prolonged period of no or slow growth and that intensive surveillance has generated the excess of melanomas among Laboratory employees remain central questions.

Issue No. 2: Can the excess be attributed to occupational exposure?

In the Austin and Reynolds' report, five occupational factors were cited as accounting for much of the observed excess in malignant melanoma among Laboratory employees. These five factors - exposure to radioactive materials, volatile photographic chemicals, Site 300, chemist duties, and Pacific Test Site, all of which were assessed for the ten years prior to the diagnosis of melanoma among cases - emerged after numerous exposure differences between cases and controls in the Laboratory population were analyzed. Many reviewers were pointedly critical of the Austin and Reynolds' report for overstating their conclusions about these five factors, for the following reasons. These factors were not explicitly identified in any prior hypotheses about their relationship with melanoma. In the absence of a priori hypothesis and following a multitude of analyses of exposure differences between cases and controls, the likelihood of finding false positive associations increases greatly. There is no prior solid experimental evidence for linking melanomas with these occupational hazards. In particular, and in spite of a large body of experimental and epidemiologic research, ionizing radiation has not been shown to induce or promote malignant melanomas. A study of melanoma incidence at the Los Alamos National Laboratory produced no evidence of excess incidence among workers at that Laboratory. Further, while handling of radioactive materials at the Lawrence Livermore National Laboratory was statistically associated with increased risk of melanoma, melanoma risk was not related to radiation exposure as measured by film badge dosimeters. The latter is a direct measure of external radiation exposure, while "handling of radioactive materials" is a surrogate for radiation exposure.

Overall, the conclusion that the five cited occupational factors are causally related to the excess of melanoma is unwarranted. The Austin and Reynolds' study was an exploratory analysis and was not designed to test an hypothesis about specific occupational causes of melanoma. Chance findings are highly likely in such exploratory studies, and the likelihood of chance findings increases with the multitude of factors evaluated. The relationship of these five factors with risk of melanoma lacks biological plausibility, shows no biological gradient, and lacks both internal consistency (e.g. no association with measured radiation dose) and external consistency (no confirmatory studies in other exposed populations or in animal studies).

Issue No. 3: Non-occupational factors

Given the small number of melanoma cases (31), several reviewers expressed concern about the ability to identify the independent effects of occupational and non-occupational risk factors on melanoma risk. To do so requires partitioning the 31 cases into multiple strata of both occupational and non-occupational variables; this causes considerable instability in relative risk estimates for the effect of one factor adjusted for several other factors. In general, the Austin and Reynolds' study replicated previous findings regarding several non-occupational risk factors, namely ease of sunburning, presence of numerous large moles, and having an advanced educational degree. Weak associations were found with hair and eye color, unlike other studies. Austin and Reynolds were criticized by some reviewers for their method of assessing eye and hair color, and for limiting their history of outdoor exposure to years after age 21, since exposure during juvenile years might be critical in affecting melanoma risk. In general, however, the analysis of non-occupational factors was not a major point of contention in the reviewers' comments.

Issue No. 4: Were appropriate methods and analyses used?

The case-control approach to the study of melanoma risk factors in the Laboratory population was efficient and appropriate. By drawing both cases and controls from the same population (all were Laboratory employees), the investigators avoided a common form of selection bias that occurs when controls are drawn from a population (e.g. hospitals) unrepresentative of the exposures of the population from which cases arose. Also, by making exposure estimates based on both objective personnel records and on personal interviews the investigators were able to utilize information on life-style and personal as well as occupational factors.

The major concerns in the reviewers' comments under this issue were the small sample size, the multitude of comparisons in the absence of an explicit prior hypothesis, and the consequent overstatement by Austin and Reynolds of the causal relationship between occupational factors and melanoma. These points were discussed under Issue No. 2, above. In addition, one reviewer comments that the failure to evaluate the goodness of fit of the logistic regression model, which dominates the analysis in this study, leaves the reader uncertain about the appropriateness of this model.

Conclusions

The general consensus of the seven reviewers is that occupational exposures at Lawrence Livermore National Laboratory have not been established as a causal factor for the observed excess of malignant melanoma. Several observations support the impression that some or all of the observed melanoma excess may be attributable to intense surveillance and enhanced detection of early stage melanoma lesions. Since the incidence of melanomas among Laboratory employees has not diminished, an early harvesting effect is unlikely. This suggests the distinct possibility that localized, in situ melanomas that would normally not be detected are being reported, and that in the absence of this enhanced detection, many of these early stage lesions would show little or no clinical progression. This phenomenon would explain the continued high incidence of melanomas in the absence of a physical or chemical inciting cause. A key point in this reasoning is the issue of the rate of growth of early stage melanomas, and this point remains a key question for study.

Even if the observed excess cannot be explained by detection bias, the reviewers agree that the Austin and Reynolds' study does not make a convincing case for occupational factors being a cause of the high melanoma incidence. The number of cases was too small and the number of exposure factors analyzed was too great to allow acceptance of a causal hypothesis. The biological plausibility of the causal hypothesis about occupational factors is not established and evidence for a dose-response gradient was not provided. The relationship of melanoma with radiation exposure is neither internally or externally consistent.

A. Introduction

Since 1972, the incidence of malignant melanoma (MM) among employees of the Lawrence Livermore National Laboratory (LLNL) has been observed at three to four times the rate in the surrounding community. From the beginning of the Laboratory in 1952 through 1971, the incidence of MM was either zero or one case per year. This annual incidence was less than or equal to that expected (Moore et al., 1984) based on incidence rates reported to the tumor registry encompassing Alameda County, where 80% of LLNL employees reside. Between 1972 and 1976, 14 new cases of MM were found at LLNL; at this time the medical services department became sufficiently concerned to initiate some studies of MM incidence; it appeared that twice as many MM cases had occurred as would be expected. In 1977, Dr. Donald Austin, Director of the Resource for Cancer Epidemiology Section in the California Department of Health Services and head of the Northern California Tumor Registry encompassing five counties in the San Francisco Bay Area, was commissioned by LLNL to conduct an in-depth epidemiologic study of MM risk among LLNL employees. This emanated in several reports. Report No. 1 of April 1980 compared the incidence rate for malignant melanoma among LLNL employees with the surrounding community and reported a 3-4 fold excess at LLNL. Report No. 2 of February 1983 showed that the incidence of other cancers was not in excess at LLNL. The most comprehensive study was initiated in 1980 and emanated in Report No. 3 entitled "A Case-Control Study of Malignant Melanoma Among Lawrence Livermore National Laboratory Employees", and was authored by Donald F. Austin and Peggy Reynolds, dated July 3, 1984 (unpublished to date). This report is subsequently referred to as the Austin and Reynolds Report #3.

Because of the importance of the report to LLNL management and employees, and in part due to the controversial nature of some of the report's conclusions, seven health scientists, including biostatisticians, clinicians specializing in dermatological diseases, and epidemiologists were asked by the Melanoma Investigation Task Group formed within LLNL to review the Austin and Reynolds Report #3 in terms of its substantive findings and the epidemiologic and statistical methods used to derive its conclusions. This request was made in September, 1984 and replies were received in the next three months.

The Melanoma Investigation Task Group at LLNL decided to commission an epidemiologist external to LLNL and to the Department of Energy to synthesize the comments of the above seven reviewers and to draw conclusions based on these written comments. The authors of the present synthesis are epidemiologists specializing in occupational and environmental

epidemiology, with emphasis on cancer risks associated with occupational exposures. A major purpose of commissioning a synthesis by an external group was to avoid the appearance of conflict of interest on the part of the LLNL management. This has been carefully pursued, since at no time did any LLNL employee suggest to the present authors any particular emphasis or direction to be taken in preparing this report. Although our conclusions are our own, we found it easy to base our conclusions solidly on the direct statements of the seven reviewers. There was no major issue on which the reviewers disagreed among each other.

In preparing this synthesis, we were requested to use direct quotations from the seven reviewers' comments. However some of the reviewers preferred not to be quoted by name, and we have elected to assign code letters, A through G, to their comments. A list of the seven reviewers is appended to this report; code letters do not correspond to first or last names of the reviewers.

The comments of the seven reviewers were organized by us around four major issues emerging from the Austin and Reynolds Report #3. These issues are:

(1) Is there a real excess of malignant melanoma at LLNL?

(2) If there is an excess of malignant melanoma at LLNL, can this excess be attributed to occupational exposures?

(3) Were non occupational factors adequately addressed?

(4) Were appropriate methods and analyses used?

In the following section (Section B) of this report, we address each of these issues in turn. In doing so, we cited appropriate comments by the seven reviewers and attempted to limit our commentary within the scope of their comments. In Section C we present our conclusions, which we believe are in close agreement with the overall tenor of the reviewers' comments. In Appendix A, we provide a concordance of the comments by each reviewer on each of the four major issues cited above. This concordance is intended to let the reader be exposed to direct quotations from all seven reviewers as their comments bear on the four issues.

B. Discussion of the Four Major Issues

Issue No. 1: Is there a real excess of malignant melanoma at LLNL?

The reported number of LLNL employees with MM is significantly greater than expected, when expected numbers are based on the frequency of incident MM lesions reported to the five-county population-based tumor registry of Northern California. The reality and statistical significance of this excess is not challenged by any of the reviewers. However, four of the seven reviewers (Reviewers A,B,D,F) state that the observed MM excess may be due to detection (or "surveillance") bias, brought on by the reaction to the initial report of a cluster of MM cases in LLNL employees between 1972 and 1976. This reaction may have led to intensive surveillance and vigorous diagnostic evaluation of suspicious moles among LLNL employees from 1977 to the present, thereby producing a real increase in the observed incidence of MM when compared with the incidence in the general population of the region. But this increase may not be attributed to exposure differences between LLNL employees and the general population.

The arguments for detection bias need to be considered in a historical perspective. In 1976, the Chief of Medical Services at LLNL as well as several physicians in the Livermore area independently observed an unusual cluster of MM cases among Laboratory employees. Figure 1 (p.115) of Report #3 by Austin and Reynolds (1984) enumerate MM cases by year from 1960 to 1984, as follows:

Year	# Cases	Year	# Cases
1960	1	1972	4
1961	0	1973	0
1962	0	1974	4
1963	1	1975	4
1964	1	1976	2
1965	0	1977	6
1966	0	1978	1
1967	0	1979	3
1968	1	1980	8
1969	1	1981	2
1970	1	1982	1
1971	1	1983	4

From 1960-1971, the annual MM incidence rate at LLNL was the same as the rate in Alameda County (Moore et al., 1984). Then from 1972-1976, the MM incidence rate rose to three times the Alameda County rate and was probably responsible for the initial

observation of a MM cluster at LLNL. These observations were widely publicized at LLNL and in local newspapers in 1977 and thereafter. The spike in MM cases in 1977 could be the result of greatly intensified surveillance by local physicians and by concern among employees. The study of the excess MM incidence at LLNL was completed by Austin in 1980, confirming the high incidence of MM (Austin, 1980). As reported by Moore et al. (1984) the announcement of these results was attended by much publicity, and this may again have stimulated intense concern and medical surveillance, possibly accounting for the second spike of cases in 1980.

Hence it is possible that the cluster of cases in 1972-76, which itself may have been a chance deviation in the incidence of MM at the Laboratory, may have generated greatly enhanced detection of early MM cases in the years 1977-83.

Austin's arguments against detection bias brought on by differences in medical ascertainment, are given in pages 43 to 47 of his Report #3. Briefly, these arguments are:

(1) "The magnitude of the MM excess among LLNL employees during 1972-1977 was 3 to 4 times that for the age/race/sex/geographic matched segment of the population in the Bay area" (p.43 of Report #3). These cases preceded any awareness of the existence of a MM excess among LLNL employees.

(2) "If the MM excess were a function of generally enhanced medical surveillance one might expect that the overall incidence of all cancer, particularly in situ cancers, in this employee population would be likewise artificially elevated. This is not the case...." (p.44, Report #3).

(3) "The second alternative for better surveillance to create the 3- to 4-fold excess among LLNL employees (when no real increase in MM is occurring) is a biopsy rate of pigmented lesions at 3 to 4 times the rate of non-employees. If the increased biopsy rate were not due to an actual increased disease rate, then it could result in an apparent increased incidence in two ways. Either the cases destined to be diagnosed at a later date are diagnosed earlier... or a large number of nonmalignant lesions are incorrectly diagnosed." (p.44, Report #3).

(4) If increased surveillance of MM prompted earlier diagnosis and treatment, "one would expect a) a decline in subsequent observed incidence, and b)

LLNL cases to present in earlier stages of disease than those in the general population." (P.44-45, report #3). However MM incidence has not declined subsequent to 1977. While there is evidence to suggest that more recent cases are being diagnosed earlier, this trend appears to be similar to that observed in other case series (p.45, Report #3).

(5) "One may consider the possibility that the earliest staged invasive MM cases... may represent cases which, if not diagnosed and treated, would regress and never present as a more advanced lesion. Such an argument is highly speculative.... The diagnosis of early MM... carries a low but real increased mortality from disease.... In addition, MM's evidencing areas of regression have a worse prognosis than those without areas of regression..." (pp.46-47 Report #3).

Each of Austin's five arguments received some discussion in the comments of Reviewers A,B,D,F.

Argument No. 1: the 1972-76 excess

Reviewer A states that chance cannot be ruled out and that Austin's conclusion on page 58 (conclusion #1) which categorically accepts the cluster of cases as a "real" excess is overstated. Reviewer B suggests that the general quality of medical services provided by the LLNL Medical Department and Kaiser Health Plan was superior to that of the comparison population and could have been responsible for better case ascertainment at LLNL, even prior to the recognition of the cluster of MM cases in 1976. Reviewer D points out that Austin's comparison of LLNL incidence rates with rates in the local population may be biased by the fact that the MM rates in the local population are based on cases reported through hospital tumor boards and pathology departments. While this provides complete case enumeration for internal malignancies, this method of case finding may well be inadequate for malignant melanoma which is sometimes diagnosed and treated in clinics by dermatologists without hospitalization and sometimes without review by a pathologist.

Reviewer E considers that the case is made for a real elevation in risk of MM among LLNL employees.

Other reviewers do not directly address this argument.
In summary, several of the reviewers are skeptical that chance can be ruled out as explaining even the 1972-1976 excess.

Clustering of cancer cases at one place is not an uncommon phenomenon, in the absence of any identified causal factor. Furthermore, the possibility of a detection bias explaining the 1972-1977 excess is also raised, and could be accounted for by better medical care for employees at LLNL and more complete melanoma case enumeration at LLNL compared with the surrounding population.

Argument No. 2: In-situ cases at other sites should be increased.

While not specifically addressed by any of the reviewers, this argument makes little sense. Once the 1972-1977 excess was reported, the focus of concern would naturally be on malignant melanoma, not on prostate, cervical or breast cancer, sites for which careful cancer screening might lead to discovery of in situ cancers. There is nothing to suggest that the 1972-76 melanoma cluster generated intense surveillance for cancers at all sites; this generalized screening would be costly, whereas inspection for moles can be accomplished by the subjects themselves at no monetary cost. Concern for a high melanoma risk is unrelated to cancer risks, other than skin cancer.

Argument No. 3: A high biopsy rate

Reviewer B comments on the higher level of education among LLNL employees than in the general population, and notes that this educational level would be likely to increase the degree of sensitization of the employees to the risk of melanoma subsequent to the 1976 recognition of an excess melanoma incidence. The reviewer says: "These characteristics, and possibly others, might have led them to self examination and to seek medical examination for melanoma more than was occurring in the general population." Later the reviewer speculates that "biopsy and pathologic examination of tissue may have been used more extensively. Legal considerations that evolved may have enhanced the tendency."

Reviewer D, commenting on the Hiatt and Fireman report of February, 1984, points out that these investigators observed that LLNL employees have significantly more skin biopsies than the general population.

Reviewer F notes that "among Kaiser Health Plan members, employees of the LLNL had more biopsies for skin lesions than their matched controls from Walnut Creek." This reviewer offers three explanations for this finding: more melanomas are being harvested; there are more suspicious lesions requiring biopsy; both harvesting and a true increased incidence of melanoma are

occurring. The reviewer tends to discount early harvesting of lesions which would later be diagnosed, because melanoma incidence at LLNL has not returned to levels reported in the surrounding population. However, he states that the "possibility that some lesions are being picked up that would never be diagnosed is, I believe, a matter of debate and not 'rather remote'", as Austin claims. Reviewer F continues: "There is no data on humans which says that malignant melanoma completely regress," though there is "indirect evidence which can be brought to bear... First, it is the natural history of pigmented nevi to regress. They appear in young adulthood, progress to dermal and some time pedunculated lesions and then disappear in later life... Also, spontaneous regression of primary melanoma is known to occur even though it is extremely difficult to document." Hence, this reviewer believes that a harvesting of melanomas that would otherwise have regressed cannot be eliminated from consideration.

The arguments of Reviewer F would seem to counter Austin's assertion (page 44) that for the high biopsy rate among LLNL employees to occur in the absence of a true increase in incidence of MM, a large number of nonmalignant lesions would have to be incorrectly diagnosed. All, or nearly all, of the melanomas diagnosed in LLNL employees may be correctly diagnosed, but the diagnosis does not predict the subsequent progression or regression of these lesions. Reviewers B, D and F are suggesting that enhanced concern may have led to a high biopsy rate. A number of these biopsied lesions diagnosed as melanoma may have remained static or progressed very slowly in the natural course of events. The fact of the high biopsy rate seems to be established. The issue of the rate of progression of these melanomas is unresolved and possibly unresolvable, since, as Reviewer F points out, diagnosis of melanoma requires biopsy and biopsy often removes the entire lesion, making it impossible to study its progression or regression.

Argument No. 4: (a) Expect a temporal decline in melanoma, and (b) an earlier stage of melanoma.

Increased surveillance would not be followed by a decline in the observed incidence of melanoma after 1977 or even after 1980 if heightened awareness resulted in detection of lesions that would not otherwise be diagnosed during the study period and if this intensified surveillance persisted beyond 1980. Reviewer F says: "I would predict that the current intensified surveillance program will result in a further increased incidence of malignant melanoma and that the incidence will remain high as long as the surveillance and awareness persist." The high biopsy rate for skin lesions among LLNL employee members of the Kaiser Health

Plan compared with other persons served by the Health Plan, as reported by Hiatt and Fireman for the years 1974-1980, supports the impression of persistent, increased surveillance of LLNL employees.

That melanoma is being diagnosed at an earlier stage among LLNL employees is shown in Table 7 (page 75) of Austin's Report #3. Between 1969 and 1976, 7 of 12 cases with known depth of invasion were classified as deep lesions (greater than 0.75 mm), while between 1977 and 1980 2 of 13 were classified as deep. When characterized by stage of invasion (Clark's microstaging levels, ranging from Level I - in situ - to Level IV, the most invasive), Table 6 (page 76 of Austin's report) shows 7 of 13 staged lesions being in the Levels III or IV category (more evidence of invasion) for melanomas among LLNL employees diagnosed between 1969 and 1976, while only 1 of 14 lesions were in these more invasive categories between 1977 and 1980. These data demonstrate a shift in the stage of detection of melanomas after 1976 towards more superficial and more localized lesions. Reviewer A notes Austin's comment that this shift is a general trend and declares "the generalizability of such academic observations are not great, and the possibility that increased surveillance at LLNL produced an apparent continued high incidence cannot be completely excluded." Reviewer B emphasizes that Austin's report "gives scant attention" to the question of whether the reported excess of melanoma reflects a tendency of the disease to be diagnosed and reported more than in comparison population groups. Reviewer B goes on to suggest a "pathological/epidemiological study" that might address this question.

Argument No. 5: Regression of early staged melanomas is highly speculative.

Reviewer F disagrees with Austin and cites evidence from the natural history of pigmented nevi, which show regression between young adulthood and later life. This reviewer also states: "Spontaneous regression of primary melanoma is known to occur even though it is extremely difficult to document." Regression is not the only alternative to progression of melanomas to more invasive lesions. It is possible that in situ lesions may remain static for prolonged periods an then progress slowly or more rapidly. The natural history of in situ melanomas is clearly an important issue in resolving the question whether intensive surveillance is merely causing an early harvest of lesions. At the present time this natural history is unknown.

In summary, among the reviewers who commented on the issue of an excess of melanoma among LLNL employees, four of five reviewers feel that chance and detection bias could account for

the excess, and that a true increase of melanoma in LLNL compared with the surrounding population is not established. Heightened awareness among a well educated and medically well served group could result in intensified surveillance and discovery of early stage, superficial melanomas, especially after 1976 when the initial cluster of cases was reported. The combination of intense surveillance and discovery of early lesions would increase the observed incidence of melanoma in the LLNL population, even in the absence of a higher level of exposure to an etiologic agent than in the general population. If these superficial and early lesions would naturally regress or remain static for prolonged periods, then in the absence of intense surveillance, an excess of melanoma would not be observed among LLNL employees during the course of the study, assuming that similar superficial and early lesions are not being reported to the tumor registry serving the surrounding population.

The paper by Moore, Bennett and Mendelsohn (1984) entitled "Melanoma Among LLNL Employees: An Epidemiologic Puzzle" contributes some useful evidence to the hypothesis that a number of the melanomas reported at LLNL are early, non-invasive or slowly progressive lesions. Figure 5 in their report shows a pattern of reduction in melanoma thickness over time among LLNL cases, a pattern not found in the University of California, San Francisco melanoma clinic. Measurements of thickness in this study were reported by the same dermathopathologist. These authors also report on mortality experience among LLNL employees from 1964 through 1979, compared with mortality expected on the basis of U.S. vital statistics. The data show that as of the end of 1979, LLNL was not experiencing a significant increase in melanoma mortality (the Standardized Mortality Ratio for skin cancer was 139, with 95% confidence limits of 51 to 302). In the face of the reported 3-4 fold increase in melanoma case incidence, these mortality data suggest a low case fatality ratio among diagnosed cases, again supporting the hypothesis that early and superficial lesions which are not progressing are being detected at LLNL. A serious limitation to the use of the mortality data to support this hypothesis is the long delay in metastases of some melanomas, delays of up to 10 to 15 years after first diagnosis. Hence the study period 1964-1979 may well not be long enough to draw conclusions about the metastatic and fatal nature of the incident melanomas discovered since 1972 at LLNL.

Whether the melanomas being detected at LLNL are often lesions likely to remain static or regress spontaneously is probably unknowable, for reasons given previously. The argument as to whether the observed excess is due to detection bias or due to exposure to an etiologic agent hinges on this point. Reviewer

B describes a "pathological/epidemiological study" that might be helpful. In our opinion this suggested study would not directly resolve the issue. More needs to be learned about the natural history of early and superficial melanomas, particularly with respect to their incidence in a population similar to LLNL and with respect to the probability of metastases of these lesions.

Issue No. 2: If there is an excess of malignant melanoma at LLNL, can the excess be attributed to occupational exposures?

a. Summary of Methods and Findings

Austin and Reynolds (Report #3) obtained Data on occupational exposure factors from three sources: 1) personnel file classification records, 2) information from respondents regarding job and task descriptions, and 3) respondents' reports of exposure to specific substances. The last two sources were obtained by means of mailed and in-person interview questionnaires; the personnel file data are independent of the questionnaire responses.

Subjects were classified from the personnel records into major occupational groups (Administrators, Scientists, Technicians and Clerks) and into subclassifications within those categories (e.g. physicists, chemists). The questionnaires elicited information on all jobs held since age 16, including types of industries, job duties, plant and laboratory locations. Detailed job description and exposure information was obtained for all jobs held during the 10 years immediately preceding each case's date of diagnosis; thus a single index date was used for each case and his matched controls.

Comparisons of cases' and controls' work and exposure histories revealed that five occupational factors at LLNL were relatively more common among cases. These were: 1) exposure to radioactive materials, 2) more than one visit to a non-nuclear weapons testing site (Site 300), 3) exposure to volatile photographic chemicals, 4) presence at a nuclear testing site in the Pacific, and 5) employment as a chemist. The investigators determined, by means of statistical analyses of these data, that each of these factors independently conferred an excess risk for MM, and that, taken together, these five factors could be responsible for the four-fold MM excess at LLNL.

Two other exposure factors emerged as being more frequent among cases than controls; exposure to high explosive fumes, and work location in a building constructed in 1969. Austin and Reynolds conclude that exposure to high explosive fumes is a secondary consequence of exposures to radioactive substances and

work at Site 300, thus this exposure is not in itself a likely causal factor. There is no clear explanation given for the association with employment in a building constructed in 1969, although the investigators note that the pattern of excess risk suggests a possible point-source epidemic (i.e., clustering in space and time), or alternatively, that whatever relevant exposure(s) exists in these buildings may promote rather than initiate the development of MM.

b. Critique of Occupational Factor Analysis and Interpretation

(1) Study Design and Statistical Analysis

The case-control design used by the investigators is a common and efficient method of identifying specific occupational factors that relate to disease risk. Austin and Reynolds made especially efficient use of this strategy insofar as data on non-occupational risk factors were obtained. Typically, data from epidemiologic studies conducted in occupational settings are restricted to occupational factors, thus the contribution of non-occupational factors in those studies is undetermined or subject to speculation. Matching of cases and controls with respect to age, race and sex is a standard procedure that is justified in most studies because of the frequent associations of these factors with disease risks. Random selection of controls from within the matching factor categories is a means of ensuring against selection of a biased comparison group, e.g. with atypical exposure characteristics, and also enhances generalizability of the study findings to the entire source population, in this case, all workers at LLNL.

The investigators' attempts to obtain reasonably thorough occupational exposure information from several sources is commendable. Obtaining such data from personal interviewing of subjects, rather than mere reliance on personnel record information, is the exception rather than the rule in occupational epidemiology. However, there is some concern about the investigators' focus on detailed occupational data only for the 10 years preceding the cases' diagnoses as noted by Reviewer E. This type of data truncation may result in an omission of pertinent exposure data (i.e., before the preceding 10 years), especially in view of the commonly accepted long latency period between exposure onset and the manifestation of most cancers. On the other hand, the restriction of detailed data to this 10-year interval may serve to avoid the introduction of inaccuracy caused by vague recollections of far distant exposures.

The statistical analysis methods used were simple odds ratio calculations and multiple logistic regression modeling techniques. The odds ratio is computed in case-control studies

as an estimate of the relative risk, i.e., the rate of disease among persons with some exposure characteristic divided by the corresponding disease rate in persons without that characteristic. Odds ratios were computed separately for each of the exposure factors. Multiple logistic regression analysis is a technique that is used to estimate relative risks associated with the exposure factors. The difference between the two methods is that simple odds ratios estimate relative risks for exposures without taking account of other exposures and other potentially relevant factors. By contrast, logistic regression analysis facilitates estimation of the effects (relative risks) for each factor above and beyond the effects of other exposures and co-variables (e.g. age).

Both analytic techniques have particular advantages and limitations. The simple odds ratio method has a clear interpretation, but may yield misleading results for individual exposure factors if the effects of several factors are highly interrelated or dependent on one another. Thus, for example, an odds ratio of 5.0 may be observed for factor A, and an odds ratio of 3.0 may be observed for factor B, which superficially indicates that factor A is the more potent cause of disease. If, however, many of the subjects exposed to factor A were also exposed to factor B, then the two results would not be independent. Independence of effects for factors A and B would be assessed by determining the relative risks associated with A among persons without B, and vice versa. Logistic regression modeling was developed and is used for the purpose of isolating independent as well as interactive effects of multiple exposures. The data in Table 39 (p.109) of the Austin and Reynolds report illustrate the use of logistic regression analysis. Here, the relative risks are presented for each of the identified important occupational factors, while taking account of (controlling for) the effects of non-occupational factors. (The findings from this important table will be discussed in more detail subsequently.)

While the statistical analysis methods are appropriate for case-control data, several concerns regarding the interpretation of the findings can be raised. The first, and most influential, criticisms pertain to the small sample size of study subjects and the large number of factors examined in this study. With a relatively small number of cases (31) it is to be expected that the relative risks are likely to be subject to substantial numerical instability. An example of this circumstance is the relative risk for chemist duties estimated at roughly 8 (Table 23, p.92), which was based on only 4 chemists in the case group. Had there been only 3 chemists among the cases, then the relative risk would have diminished considerably.

As a means of evaluating the stability of the results, Austin and Reynolds computed probability values (p-values) for each factor. Ostensibly, these p-values indicate whether or not relative risks of the magnitude seen are statistically likely to be mere chance observations. In a strict sense, a p-value has meaning only when one is testing a specified a priori hypothesis. While their choice of study factors in a general sense implies certain prior hypotheses (and hunches), this study is most realistically viewed as exploratory or hypothesis generating rather than hypothesis testing. This is not to say that exploratory studies are necessarily less valuable than hypothesis testing investigations. In fact, given the paucity of knowledge regarding MM etiology and the unexpectedly excessive frequency of MM at LLNL, a hypothesis generating study is certainly warranted.

The investigators' choice to place a considerable emphasis on statistical significance is a point of contention. As Reviewer C points out, there is on average a 1 in 20 chance (corresponding to the usual 0.05 nominal p-value level for deciding statistical significance) that 1 of 20 comparisons will appear to yield a significant result strictly by chance. This reviewer cites the two statistically significant findings (Site 300 and Pacific Test Site) out of the 35 case-control comparisons in Tables 26 and 27 combined as an illustration of this point. Reviewer F similarly remarks on the potential for spurious chance findings, given the absence of prior hypotheses and the use of multiple statistical significance tests. Austin and Reynolds are cognizant of this potential source of spuriousness, yet statistical significance testing is used repeatedly as a guideline for interpreting findings and invoking causality. There are well known statistical methods to correct, or adjust, observed p-values for the number of tests performed, yet the authors prefer to present unmodified p-values. However, such adjustment techniques are seldom used except for purposes of stringent hypothesis testing, and even these methods do not ensure non-random findings.

The smallness of the sample size cannot be overcome even with logistic regression modeling, for here again, relative risk results may still be unstable (Reviewer E). Moreover, an inadequate mathematical model resulting from sparse data on both non-occupational and occupational factors is likely to compound the uncertainty in the estimates of the exposure effects (Reviewer G). Reviewer G further comments on the investigators' apparent failure to test the appropriateness of the logistic models developed in the analysis. There are formal mathematical methods to evaluate how well a logistic model actually fits the data. The implication here is that the assumed (i.e., logistic) models may not conform to the true underlying mathematical

function which describes the relationship between the exposure factors and MM risk. The logistic model (as is true of any mathematical model) has a specific theoretical form which is a represented by a sigmoidal relationship between exposure and risk. It is possible to impose a logistic model to case-control data as a means of estimating relative risks; however, when the data do not conform to a true logistic form, because of unusual distributional properties or small number instability, the results obtained from the analysis will be of questionable validity. Given the importance placed on logistic regression results (viz. Tables 33-41), there should have been some mention of tests for model fitting.

(2) Identified Occupational Risk Factors

(a) Radioactive Materials

Ionizing radiation is a recognized human carcinogen that affects multiple organs sites and tissue types. Consequently, radiation is a logical candidate as a suspect risk factor in any study where occupational radiation exposures occur. Austin and Reynolds observed an apparent association of MM with the handling of radioactive materials; roughly 65 percent and 33 percent of the cases and controls, respectively, reported such exposure, and the adjusted odds ratio from logistic regression analysis is nearly 6 (Table 39, p.109). Radioactive materials handling shows the strongest association of any of the occupational factors when non-occupational exposures are taken into account (Table 39).

The importance of this observed association is questioned for several reasons, apart from the previously mentioned issues of multiple testing and numerical instability. First, there is scant evidence in the scientific literature to support an etiologic relationship between ionizing radiation (not ultraviolet radiation which is a form of non-ionizing radiation, and is a well established MM risk factor) and MM. This point is raised by several of the reviewers, particularly Reviewers A and C. The study of MM among Los Alamos National Laboratory (LANL) workers by Acquavella et al (1982) demonstrated no effect of either external (beta, gamma, neutron) or internal (plutonium) sources of radiation. The LANL study included even fewer cases (20) than the LLNL study; nonetheless, the similarities of the two worksites makes comparisons of findings relevant. Reviewer C further suggests that the association seen in the present study may be an artifact of selective recall of exposure by cases, attributable to the association of non-MM skin cancer with MM and the heightened awareness of cases with both conditions to the carcinogenic effects of ionizing radiation. This argument is speculative, however.

A second argument against a causal role for radioactive materials exposure is the inconsistency of results. Austin and Reynolds report no association with radiation film badge dosimetry (these data are presented in some detail in the manuscript by Moore et al, 1984). It is possible, however, that film badge dosimetry may be a poor indicator of radiation exposure if the sources are internally deposited radionuclides which are weak gamma or beta radiation emitters. Austin and Reynolds do not describe exactly what the radioactive materials were, which hinders interpretation of the findings. The absence of excessive rates of the commonly recognized radiation-related cancers (lung, leukemia, non-Hodgkin's lymphoma, thyroid, bone) is discussed in the report and by one reviewer (Reviewer F) as an example of inconsistency. One would not necessarily expect excesses of all or even most of the radiation-related cancers, yet the apparently pronounced effect for MM, a cancer seldom linked to ionizing radiation, suggests an anomalous finding.

Finally, as Reviewer A discusses, there is no evidence of a dose-response relationship between radiation and MM risk in this study. Table 30 (p.99) reveals that the majority of the association is accounted for by an ever vs. never comparison of exposure, and that length of exposure and total exposure are virtually unrelated to MM risk. Dose-response relationships, where the risk increases as exposure intensity and duration increase, strengthen causal arguments. There is a voluminous body of scientific literature on radiobiologic effects which clearly documents dose-response relationships for many types of ionizing radiation and human cancer incidence. The absence of a dose-response trend in this study diminishes the importance of this observation. One possible, but not necessarily likely, alternative causal explanation might be that brief, intense exposures to radioactive materials may have been etiologically significant. A better description of the nature of radiation exposures at the LLNL worksites is needed to address this concern.

(b) Site 300 Assignment

The exposures of relevance presumably are related to non-nuclear weapons testing. No specification of the suspect agents is presented, thus this association has limited meaning, and may be a statistical artifact.

(c) Exposure to Volatile Photographic Chemicals

Approximately 35 percent of MM cases reported inhalation exposures to volatile photographic chemicals as compared with 15 percent of controls (Table 28, p.47), thus resulting in a

relative risk of about 3 for this factor. By contrast, the relative risk associated with potential skin exposures was about 2. If the exposures resulting from skin and inhalation exposure routes are the same, then the differential effect suggests lack of consistency. In fact, it would be expected that dermal exposure should be more relevant than inhalation exposure to the same agent. It should be recognized, however, that inhalation exposures were to presumed "volatile" chemicals which may differ from chemicals encountered dermally. Austin and Reynolds do not indicate which specific agents are implicated.

The risk associated with volatile chemicals appears to be dose-related; that is, risk increases as a function of increasing intensity and duration of exposure. This result is demonstrated in the data of Table 30 (p.99). The authors' interpretation of the pattern of the relationship between volatile photographic chemicals and MM (p.25) is that risk is probably increased each time an exposure occurred, and that exposures probably occurred routinely (possibly, every day). Austin and Reynolds appropriately point out the crudeness of this dose-response assessment technique, and caution against overinterpretation of the trends.

No association was seen with photographic chemicals encountered during hobbies as noted by Reviewer A. This absence of an effect doesn't necessarily weaken the association with occupational exposure since hobby exposures may differ from occupational exposures with regard to types of chemicals used and intensity of exposure.

(d) Pacific Test Site Assignment

Presence at a nuclear testing site in the Pacific at the time of a nuclear test was reported by 13 percent of cases and 4 percent of controls, resulting in a relative risk of roughly 4. The investigators argue for an effect of this exposure, which is independent of handling radioactive materials, although the nature of this factor suggests ionizing radiation as the exposure of relevance. As such, the above comments pertaining to radiation and MM are applicable here.

(e) Employment as a Chemist

There were only 4 cases and 2 controls who reported employment as a chemist; the apparent relative risk is approximately 8 or 9 (depending on the method of analysis). Despite the small number of subjects who were chemists (by their personal accounts rather than from job classification records), Austin and Reynolds note that this factor confers the largest individual risk for MM.

There are reports in the literature of elevated MM risks among chemists, oil refining workers and PCB-exposed workers; however, the evidence is conflicting, and with the possible exception of PCBs, no specific chemical risk factors have been identified (Reviewers C and F). No association of MM with employment as a chemist was observed in the Los Alamos National Laboratory study (Acquavella et al., 1983).

The relationship of MM risk with chemist employment in this study does not implicate any specific hazardous substance. In fact, as Reviewer A notes, there is general inconsistency with regard to occupational chemical exposures, as evidenced by the absence of associations with building location, skin exposures and use of protective equipment. Furthermore, the absence of an association with employment as technicians, some of whom share common workplace exposures as chemists, suggests to Reviewer E that the risk among chemists may be spurious. One possible explanation for the difference in effect for chemists and technicians may be a higher level of education among chemists. MM has frequently been related to higher socio-economic and educational status, and in this study advanced education is strongly associated with MM (RR = 3.4, Table 7, p.76).

(f) Exposure to Fumes from High Explosives

Inhalation exposures to fumes from stored high explosives were reported more frequently by cases (19 percent) than controls (7 percent), and the relative risk estimate obtained was roughly 3 (Table 28, p.97). The effect of this exposure persists after controlling for constitutional factors; however, there appears to be a correlation between high explosive fumes and radioactive materials exposures; this correlation causes the effect of fumes to diminish sharply when radioactive materials exposures are considered first (Table 35, p.105). The authors' interpretation here is that the association with high explosive fumes is mediated through the effects of other risk factors, in particular radioactive materials. The chemical nature of these fumes is not described, so again, specific substances cannot be isolated as risk factors. However, there is so little known about chemical inducers of MM that biological plausibility cannot be discounted.

(g) Work Location in Building Constructed in 1969

Two buildings, Experimental Physics (111) and Biomed (361) were identified as conferring an apparent excess risk. The effect of work assignment in these buildings was greatest 5 to 6 years before cases' dates of diagnoses (Table 24, p.93), which Austin and Reynolds interpret as suggestive of a possible "point-

source epidemic". Reviewer E discounts the likelihood of a meaningful case clustering by pointing out that there is a similar, parallel temporal course of work location in the "low risk" buildings constructed in 1966.

The association with the 1969 constructed buildings disappears when the cases' and controls' distributions of multiple moles is taken into consideration, despite an apparent persistence of association when the other occupational exposure factors are considered simultaneously.

There are no industrial hygiene or health physics monitoring data to characterize exposures in these buildings, and appropriately, the authors do not posit direct causality. Reviewer E's view that the treatment of a building constructed in 1969 as an indicator of exposure is "artificial and has little logical appeal" is a reasonable assessment.

(3) Interpretation of Causality

Austin and Reynolds devote a considerable portion of the discussion section of the report to issues of causality. They discuss occupational, educational and constitutional risk factors in reference to a standard set of causality criteria: absence of bias, in particular bias due to confounding; the temporal relationship of exposure and disease; specificity; strength and consistency of associations; and biological plausibility. This discussion is certainly an appropriate exercise for any epidemiologic study. Some of these topics have been addressed earlier in the reviews of individual study findings, and will not be reiterated.

What is most important when interpreting the occupational risk factor analysis results is the determination of the amount of the MM excess at LLNL that can be attributed to occupational exposures, taken singly and in combination. (For the purpose at hand, it can be assumed that the 4-fold MM excess at LLNL is real and not an artifact of diagnostic bias. This issue was discussed at length in the previous section of this commentary).

There remains the possibility that the associations observed for chemist duties, radioactive materials, volatile photographic materials, assignment at Site 300 and assignment at a nuclear test in the Pacific are merely chance occurrences resulting from an analysis of a great many factors. Statistical significance testing offers little guidance in establishing real from spurious relationships because a priori hypotheses were not tested, and because the small size of the study population may have resulted in widely fluctuating findings. These concerns have been raised

by several of the reviewers particularly Reviewers C, E and F, as mentioned earlier. Further study of MM at LLNL and at similar facilities will be informative in this regard.

Notwithstanding the issue of real or spurious findings, the authors' claims that the five identified occupational risk factors confer MM risks independently of one another and that, taken together, these five factors can account for the MM excess at LLNL warrant scrutiny. These arguments are discussed in turn.

Independence of risk factor effects means simply that a given factor can confer disease risk without requiring the presence of another factor(s). For example, asbestos can induce lung cancer even among non-smokers. The data which are pertinent to the independence of these effects are contained in Table 35 (p.105) and 39 (p.109) of the report. Table 35 shows relative risks for individual factors after adjustment is made for the other factors. It is important to recognize that each entry in this table represents the relative risk associated with a given factor, while controlling for the effects of only one other factor. Thus, the entry of 10.6 for Chemist Duties (row 9) in column 6 is the relative risk associated with chemist duties, taking into account the effect of fumes from explosives. In a strict sense, an independent effect would be demonstrated if the effect persisted while controlling for the effects of all other measured risk factors. The investigators' failure to conduct an analysis with simultaneous control of multiple factors in this table is probably a result of the small sample size. That is, mathematical models become very unstable (or incalculable) when effects from numerous factors are being estimated from a limited number of data observations. Therefore, while Austin and Reynolds may have exercised reasonable judgment by not attempting elaborate statistical modeling of the data, they have not necessarily demonstrated true independence of effects.

The data in Table 39 indicate that the effects of occupational exposures are statistically independent of constitutional and educational factors, but are not necessarily independent of each other. Here, each occupational factor is evaluated in turn in the presence of a common set of educational and constitutional variables, but not in the presence of the other occupational factors.

The problem of establishing independent exposure effects would be simplified greatly if the investigators had shown the joint distributions of occupational factors for the cases and controls, as noted by Reviewer C. In this way the reader would be able to see whether cases who were chemists also handled radioactive materials and had been assigned to Site 300, and so forth.

Independence of effects has a significant influence on the interpretation of the data in Table 44 (p.114). These data (Table 44) form the basis of the most important argument presented in the report, which is that these five factors can account for the four-fold MM excess at LLNL. Some explanation of the derivation of these data is in order.

On p.39 the authors describe a method for estimating the percent of the excess MM rate at LLNL, relative to prevailing rates in persons not employed at the facility. The expression can be represented as:

$$RR = P(R) + (1-P),$$

where:

- RR is the rate in the entire LLNL population divided by the rate in the general population assumed to be not exposed.

- P is the proportion of the workforce exposed to a particular occupational risk factor. P is estimated from the exposure frequencies of the controls, who represent a random sample of the total workforce (the cases constitute a negligible segment of the workforce, and can thus be ignored here).

- R is the relative risk among workers exposed to the factor.

This expression is meant to convey the notion that the overall (total workforce) relative risk is a weighted sum of the excesses in persons exposed and not exposed to the factor of interest. Thus, the proportion of the population exposed, P, experiences a relative risk of R times that of the not exposed workers, while the proportion not-exposed (1-P) experiences the same rate as the general population, i.e., their relative risk is 1.0 (note that a relative risk of 1.0 is not shown as the multiplier for (1-P), although it is implied). There is a tacit assumption that the rates in the not-exposed segment of the workforce are the same as that in the general, or reference, population. This assumption is necessarily defensible only if the not-exposed workers are a random sample of the reference population. There is no direct method of determining whether or not this is true, and in this study it is likely that the not-exposed workers are different from the general population with respect to education.

Austin and Reynolds computed separate relative rates for each of the 5 occupational exposure factors, as shown in the third data column of Table 44. The relative rates are expressed as percentages. Verification of these results is straightforward. Additionally, they estimated relative rates for cumulative combinations of factors taken together; these results are given in the rightmost column. The data given are insufficient to verify these findings, as the relative risks for more than one factor taken together and the corresponding proportions of exposure among controls (LLNL prevalence) are not presented. Addition of these relative rates in the manner described is appropriate provided that the relative risks are in fact independent of one another, and do not interact (e.g. the relative risk among persons exposed to two factors is greater than the simple sum of effects of the two). Neither independence nor absence of interaction has been demonstrated convincingly.

On balance, the argument that these five occupational factors can account for the excess of MM rests on a number of unsubstantiated or unverifiable assumptions. In view of this and other important concerns regarding the role of chance and inconsistency of some associations, it would appear that the authors' assertion of causal attribution for these factors is an unwarranted overstatement (Reviewers A, E and F).

This investigation does contain some provocative findings which should be explored further. The apparent excess of MM at LLNL and the general paucity of confirmed etiologic hypotheses are compelling public health and scientific reasons to pursue additional inquiry. The recommendations to continue the case-control analysis and to incorporate a more exacting characterization scheme for occupational exposures are reasonable.

Issue No. 3: Were non-occupational factors adequately addressed?

Before any conclusions can be drawn about an association between certain characteristics of employment at LLNL and malignant melanoma (MM), it is important to evaluate the role of other, non-occupational factors that might alter these associations. If other risk factors for the development or detection of malignant melanoma tend to be associated with certain job titles or chemical and physical exposures, then these job characteristics will appear to be related to MM risk, even if they are not. The LLNL MM case-control study design included mechanisms to investigate possible confounding by non-occupational factors. The questionnaire administered to all cases and controls included items on demographic risk factors

such as education, and personal, familial, and behavioral characteristics that have previously been shown to be related to melanoma risk. In some cases, objective measurements were used, such as a reflectometer for skin color. In other cases, the study relied on cases' and controls' perceptions or recall of events, such as parental history of skin cancer and the number of large moles. Austin and Reynolds present data in Tables 8-19 on the relative risk of each non-occupational factor measured. These are presented through the use of multivariate logistic regression modeling, adjusting for matching and for all other non-occupational factors in the model. From these results, the authors selected five personal and lifestyle factors to include in the model which already contained occupational risk factors. Table 39 shows the relative risk (approximated by the odds ratio) for five occupational risk factors, adjusted for the five personal and lifestyle risk factors. In our opinion, these were appropriate analyses but are constrained by the small sample size.

(1) Sample Size

The study sample consisted of 31 cases and 110 controls. Almost all reviewers expressed concern about the ability to differentiate the importance of occupational factors from other risk factors, given the small number of cases. As summarized by Reviewer C, the small number of cases is particularly important since a number of variables show case-control differences in this study. Adjusting for a number of variables in the model with a limited number of observations decreases the accuracy of the odds ratio estimates (Reviewer G). Many personal, familial, and occupational risk factors investigated in this study are likely to be related to each other. Each factor may contribute to the risk of MM independently, but given the small number of cases in this study, it is difficult to assess accurately each independent contribution. On the other hand, each of the case-control differences found may have some degree of association with some underlying, unknown cause (Reviewer C). The multivariate logistic regression analysis is an appropriate method of analysis to try and separate several different risk factors, but, as pointed out by Reviewer E, there would be greater confidence in the adequate control of confounding if stratified analysis and/or more detailed presentation of the numbers were included.

In the initial analysis of the large number of non-occupational factors that were included in the questionnaire, each potential risk factor was addressed using a pooled odds ratio of exposure among cases and controls. In most cases, answers were dichotomized into exposed and non-exposed by collapsing from each set of analyses (demographic characteristics, personal characteristics, familial characteristics,

health history, health habits, leisure time activities, and sun exposure); significant factors were chosen to be included in the multivariate analyses. On the basis of statistical significance and somewhat subjective examination of all of the factors that showed case-control differences, six non-occupational characteristics were chosen for further analysis; 1) the presence of greater than 6 large moles, 2) having a parent with skin cancer, 3) a previous diagnosis of skin cancer, 4) sunburning even with previous exposure, 5) spending greater than the median number of days inside during adulthood, and 6) having an advanced degree. To summarize non-occupational influences, all of these characteristics were entered into a multiple logistic regression model that simultaneously controlled for all other factors in the model (Table 33). Finally, all of the above factors were included in the analyses of occupational exposures to determine the importance of occupational characteristics while controlling for non-occupational risk factors.

(2) Choice and Analysis of Non-Occupational Control Variables

This study was primarily designed to determine specific occupational factors related to employment at LLNL. An exhaustive list of non-occupational factors was also included in the questionnaire to make sure that any occupational associations found were not due to other factors, and to evaluate whether the excess found at LLNL was not due to a concentration of MM risk factors among employees. Because the potential causes of the excess were unknown, it was important to collect data on a wide range of items (Reviewer C). The study results replicated previous case-control studies of malignant melanoma showing higher risk for people with certain skin characteristics such as numerous moles, poor suntanning ability, and a history of other skin cancer. This study found the traditional positive association of melanoma risk with advanced education. As pointed out by Reviewer C, replicating previous studies attests to the good design and execution of the study. There were, however, a number of issues brought up about specific non-occupational factors that were measured. These issues are discussed below.

(a) Personal and Familial Characteristics

The study found a striking association between the presence of numerous large moles and having melanoma, yet self-measurement of the number of large moles might be error-prone (Reviewer E), and cases might notice large nevi more than controls. Reviewer G also notes that dichotomizing the number of moles into greater or less than 6 might not adequately model the effect of moles on MM risk. However, the positive association with number of moles

does replicate other studies. Skin color was measured objectively through the use of a reflectometer, but hair color was subjectively measured without the use of standard hair color charts. According to Reviewer E: "This procedure would be error prone and likely to mask the expected positive associations with fair or red hair." The authors reported only a weak association with having red/blonde hair at age 20 and developing melanoma. Eye color was measured accurately but, unexpectedly, the variable did not strongly predict case status (Reviewer G). Reviewer E suggests that incorrectly grouping blue and green eyes together might account for the lack of an association. Although those who did not tan after a few days of previous sun exposure were significantly more likely to have MM, Reviewer E is doubtful whether this is "a good measure of skin response to chronic sun exposure which is an important determinant of melanoma risk."

The new finding that cases' parents were more likely to have had skin cancer than controls' parents makes sense biologically, since parents are likely to have similar skin characteristics and susceptibility (Reviewer C). As suggested by Reviewer A, however, this finding may be due to recall bias, with cases more likely to remember or know about their parents' skin cancers than controls.

Austin and Reynolds addressed the possibility that the excess of malignant melanoma might be due to a concentration of personal or familial risk factors among LLNL employees when compared to the surrounding population. They show data from a population-based survey indicating that the prevalence of moles, a history of skin cancer, a history of parental skin cancer, and the distribution of skin types were similar among Contra Costa County residents and LLNL controls. If the excess at LLNL were due solely to a concentration of high risk people, then the LLNL controls might show a higher prevalence of risk factors than Contra Costa County residents. However, this survey is a still unpublished study (Reference 49 in Report No. 3) and cannot be adequately evaluated.

(b) Past Sunlight Exposure

The questionnaire solicited data on sunlight exposure by asking for information on the number of days spent outdoors as an adult and the number of times sunburned as a youth and as an adult. These variables were included to assess the importance of non-occupational sun exposures. Only one variable related to sun exposure predicted MM risk, and this in a somewhat opposite direction. Controls were significantly more likely than cases to have spent a large number of days outside as an adult. This is consistent with the idea that people who develop MM are fair

skinned, so they avoid spending a great deal of time outdoors. Reviewer E disputes the emphasis on adult sunlight exposure in this study. He notes that limiting the "recording of outdoor exposure periods to those accumulated after 21 years of age may have missed a critical exposure period (teenage and early adult life) in which recreational outdoor exposure is common and the future of malignant nevi may be determined." The study did include data on the number of times subjects were sunburned in youth; the pooled odds ratio for this variable was elevated but not significantly. The LLNL study also neglected to investigate other outdoor jobs prior to work at LLNL, as noted by Reviewer E.

Reviewer E proposes that there may be uncontrolled confounding in the occupational associations because of inadequate measurement of sun exposure. Reviewer C affirms this indirectly by noting that even previous studies have shown that the association between sun exposure and MM is complex, and is still poorly understood, making it difficult to accurately assess the risk of MM due to sunlight. For example, it may not be cumulative sun exposure that is important, but "particularly intense exposures or exposure in a vulnerable period such as adolescence" (Reviewer C)

(c) Socioeconomic Status

The LLNL case-control studies found, like several previous melanoma studies, evidence of increasing risk with increasing years of education. Several explanations have been put forth as to why the incidence of melanoma is higher among more highly educated, affluent groups: greater intermittent leisure sun exposure, office environments, greater frequency of screening, and nutrition have all been proposed. However, it is still unclear whether socioeconomic status and education represent a combination of factors or are simply a surrogate for some unknown, underlying cause. In this study cases were more likely to have an advanced degree than controls, and this strong effect persisted even after control for other personal and lifestyle variables. Since education is strongly associated with several job titles (such as "scientist"), controlling for education is difficult when comparing the frequency of specific occupational exposures among cases and controls. Having an advanced degree is included in some multivariate analyses, but not in others, without any explanation. In order to clarify these associations, Reviewer C asks for more complete data on the distribution of occupations among cases and controls, perhaps weighted by the previously known social distribution of melanoma risk.

In the summary results presented in Table 39, advanced degree remains significant in most models, even after controlling for other occupational and personal characteristics. The

undetermined influence that education represents may affect the interpretation of the study results in subtle ways. For example, Reviewer A points out that the greater prevalence of heavy smoking among controls and the increased consumption of elective medical care among cases does "suggest some general, non-occupational behavioral differences between cases and controls". The differences in access to screening and medical care afforded by different levels of education is also important to consider, since this might result in apparent differences in the incidence of MM among different occupational groups.

Issue No. 4: Were appropriate methods and analyses used?

Given the limited number of cases of malignant melanoma at LLNL, the choice of a case-control study as a strategy to evaluate a variety of possible risk factors was highly desirable. As Austin and Reynolds state (page 1), having earlier found a 3- to 4-fold excess of MM that seemed to be restricted to the LLNL workforce when compared with the surrounding community, the investigators subsequently undertook a case-control study "to investigate what factors may have contributed to the MM excess in this occupational group". The case-control approach is particularly useful for exploring the disease risk from an array of potentially hazardous exposures.

Several strong points about the study design and selection of variables by Austin and Reynolds are worth noting. Since their case-control study was entirely based within the cohort of LLNL employees, the source population from which controls were selected was clearly representative of the population from which cases arose. This avoids one of the major biases that can occur in case-control studies, namely a selection bias that results when controls are drawn from a source population (e.g. hospitalized patients) in which (1) there is a greater prevalence than in the general population from which cases arose of other diseases caused by the primary exposure factor, or (2) there are correlates of exposure that occur more (or less) frequently in the source population of controls than in the general population. In this study, the entire LLNL employee population is the general population which gave rise to all cases of MM, and it was from this population that controls were selected. There is no evidence that other diseases known to be caused by ultraviolet or ionizing radiation were present in excess among LLNL employees. Nor was the issue of the distribution of correlates of exposure between the general population and the controls' source population of concern since these two populations were identical.

One limitation of this nested case-control approach occurs when the primary exposure responsible for excess disease is ubiquitous and evenly distributed in the entire cohort, in which

case the study would be biased against finding the causal factor. But it is highly unlikely that a causal agent would be both ubiquitous and evenly distributed in the LLNL population. Such exposures would lead to an extremely high incidence of MM, and this is clearly not the situation, since even a 3- to 4-fold excess of MM does not yield a disease incidence rate equal to that of high incidence diseases such as cardiovascular disease.

A second strength of the study was the use of both personnel records and of personal interviews to assess work history. This method of exposure assessment reduces the likelihood of recall bias. Work histories were used to characterize various sources of potentially hazardous exposures among cases and controls. These exposures are analyzed in Tables 20-29 and include such factors as job classification, work environment, scientific specialty and building location; these factors were identified from previously recorded personnel records and are therefore not subject to differential recall between cases and controls. Information obtained from personal interviews allowed the investigators to control for several known phenotypic and life style risk factors (e.g. ease of sun burns, hair and eye color, family history) while analyzing the effect of the various occupational factors. Many occupational studies fail to obtain data on personal and life-style risk factors, even though these variables often are known determinants of the disease of interest. Hence confounding by known non-occupational risk factors is not likely to explain the findings attributed to occupation in this study.

A serious limitation of Austin and Reynolds' analysis and interpretation was noted by four of the reviewers (Reviewers C,E,F,G). In the words of Reviewer C, whose views on this limitation were most forceful: "For most of the items [i.e. independent variables or risk factors] there is no previously available data to form a hypothesis that the cases would be different from controls. Hence, for these items there is no way of knowing whether the statistical significance was an expression of the multitude of comparisons, or was real". Similarly Reviewer F states: "No specific hypothesis was stated at the onset of the study and most epidemiologists would consider it an hypothesis generating study". Later, the same reviewer comments: "It is difficult to know how many comparisons were actually made in this study, but to select out the several factors that were most strongly associated with melanoma by this process may have led the authors to place more importance on them than was justified in a strict statistical sense."

In a similar vein, Reviewer E points out that "there are only 31 cases. It is questionable, therefore, whether much weight can be placed on the multivariate analyses with their...

creation, probably, of many 'empty cells'. The small numbers may explain some comparatively large changes in odds ratio (OR) estimates from one model to another".

Reviewer G criticizes the authors for failing to test or discuss the fit of the logistic model to these data. In testing for the goodness of fit, the authors would have used one of the logistic regression equations to calculate the number of cases that would have been found in each occupational exposure category as predicted by the regression model, and compared this number with the cases observed in each category. A significant difference between observed and predicted cases would suggest an inappropriate model. The authors' heavy reliance on logistic regression to calculate odds ratios adjusted for non-occupational factors warrants this criticism by Reviewer G. The particular logistic regression model chosen for each analysis may not have been appropriate, and the absence of a goodness of fit test prevents readers from evaluating this possibility.

C. Conclusions

The majority, but by no means unanimous, opinion of the seven reviewers is that two problems remain unresolved regarding the MM evidence among LLNL: (1) whether the observed excess of MM is due to enhanced surveillance and detection of skin lesions, and (2) whether the observed excess is causally related to factors in the occupational environment of LLNL.

Several factors favor the reviewers' suggestion that some or all of the observed excess of MM is attributable to enhanced surveillance and detection. The LLNL population is well educated, has good access to skilled medical surveillance, makes use of this access, and demonstrated a surge in case finding when the problem of MM excess was highlighted. If enhanced detection were solely responsible for the observed high incidence of MM, there would be one of two consequences: (a) MM cases that would eventually occur would be diagnosed at an earlier stage, or (b) MM cases that would ordinarily never be diagnosed would be discovered. If consequence (a) were true, there should be a deficit of cases after several years of observed excess. This apparently is not the state of affairs at LLNL. If consequence (b) were true, many, in fact most, of the MM cases would be localized *in situ* tumors, which would not normally be detected and would *regress*. Since 1976, there is evidence for a pronounced shift towards early stage MM lesions among the new cases at LLNL. Whether these lesions would have remained static or regressed spontaneously in the absence of intensive surveillance is unknown. This remains a key question for study.

There is a stronger consensus among the reviewers that Austin and Reynolds' conclusions regarding a causal relationship between occupational exposures at LLNL and MM risk are overstated. The authors did not begin this study with a specific prior hypothesis about occupational factors at LLNL. They had a very limited number of cases (31), and these were partitioned into many strata and in many different ways. Hundreds of comparisons of exposure differences between cases and controls were evaluated. This greatly enhanced the probability of finding false positive relationships. In the absence of a strong prior hypothesis, no adjustments in the data set can compensate for this greatly increased chance of finding "significant relationships" between exposure and disease even if no such relationship exists. Furthermore, the absence of any experimental or prior epidemiologic evidence for induction of melanomas by ionizing radiation makes the latter association implausible. The additional absence of a dose-response relationship between ionizing radiation and MM risk further weakens the argument for a causal association.

The reported relationships of MM risk with five different occupational risk factors cited by Austin and Reynolds suffer from the same limitations discussed above. These factors were not explicitly postulated in any prior hypotheses. They emerge after a multitude of exposure possibilities were evaluated. They lack evidence for a dose-response relationship, and they have no prior supporting experimental or toxicological data showing a biological basis for their relationship with MM.

In our opinion, Austin and Reynolds have not established that there is a true excess of MM in the LLNL population. The observed excess may well be the consequence of intense surveillance of moles and detection of MM that is inconsequential in terms of subsequent risk of metastatic disease. The absence of excess mortality from MM supports this conclusion but does not invalidate the alternative possibility that delayed metastases will occur. Another population of similar socioeconomic and educational composition would have to be subjected to the same intense surveillance of moles, to address the question of detection bias. Such a study could pose some ethical problems. Continued medical follow-up of LLNL employees with diagnosed MM will provide needed information about the subsequent risk of metastases associated with very early stages of MM. This follow-up will benefit not only the employees but society in general.

We also conclude that a causal relationship between occupational exposures at LLNL and risk of MM has not been established and that Austin and Reynolds have overstated their conclusions in this regard. The play of chance was not

controlled in their analyses of occupational factors. The lack of biological evidence for the reported relationships supports our rejection of a causal link. We believe that none of the reviewers disagrees with us on this latter point.

It will be difficult to conduct a follow-on study among LLNL employees to test hypotheses about occupational determinants of MM. Statistical independence of data would require elimination from study of all cases used in the Austin and Reynolds report, otherwise a new hypothesis would be tested with data used to generate the original hypothesis. Furthermore, a reasonably large sample of MM cases (say 50 or more) is necessary to test a specific exposure hypothesis, and it will likely require years of observation before this sample size is obtainable. If the issue of detection bias is somehow resolved, and the validity of the excess is established, it is probably necessary to evaluate the occupational exposure/MM relationship in other populations having a similar work environment. Such do exist, particularly with respect to ionizing radiation and chemical exposures. Thus these occupational groups may be an appropriate population to test the occupational exposure hypotheses.

It is desirable to perform further studies of the LLNL population. Exposure to ionizing radiation and to other chemicals could be characterized more quantitatively and precisely. The issue of progression of in situ lesions needs to be addressed. Individuals with in situ lesions should be followed for evidence of clinical progression either in the form of localized spread or new primary lesions.

APPENDIX A

Concordance of Reviewers' Comments by Major Issue

Comments of the seven reviewers have been organized under each of the four major issues addressed in this report. Since some of the reviewers preferred not to be quoted directly by name, we have assigned a letter code from A to G to each reviewer. The following are direct quotes from each reviewer.

Issue 1: Is there a real excess of malignant melanoma at LLNL?

Reviewer A:

"The complete characterization of cases by stage, pathologic criteria, and outcome seems to have been well done and appropriate. It does seem to me that is is impossible to completely eliminate the possibility that all or part of the excess in cases might be explained by an increase in the intensity of surveillance. The steps taken to examine this were complete and appropriate, however, and the only finding to suggest that such an explanation might be the case is the shift toward "early/thin" (Tables 6-7) melanomas which occurred as time elapsed. While the authors cite one paper suggesting that this is a general trend, the generalizability of such academic observations are not great, and the possibility that increased surveillance at LLNL produced an apparent continued high incidence cannot completely be excluded. Neither can chance be excluded as a contributor to the observation. The above notwithstanding, the continued high incidence requires a prudent observer to presume that chance is not responsible for the cluster".

"I can only repeat that neither chance nor ascertainment bias can yet be ruled out; accordingly, I also think that conclusion #3 and conclusion #1 on page 58 are overstated, the latter at least in the first sentence which categorically accepts the cluster as 'real'. Again, if it is real, neither the authors nor anyone else has a good explanation for the secular trend".

Reviewer B:

"The 3 July 1984 report by Austin continues to indicate that two questions remain unresolved:

(1) Does the reported excess of melanoma reflect a genuine excess of the disease among LLNL employees, or does the reported excess reflect a tendency for the disease among them to be diagnosed and then reported to a greater extent than in comparison population groups.

(2) If there is a genuine excess of the disease among LLNL employees, what is the cause?"

"Almost the whole of the 3 July 1984 report is devoted to the second question. The report provides highly useful data and analysis pertaining to the second question. The text gives scant attention to the first question, although Tables 4-7 emphasize its importance. The latter reveal a strong tendency toward diagnosis and reporting of less advanced lesions. Does that tendency, apparently exceeding the secular trend, account substantially for the reported excess? The report's references to cases extending beyond 1980 are helpful in regard to that question, but not sufficient."

"I still believe that before one becomes completely absorbed in the second question, it is necessary to answer the first question. Twice before I have suggested a kind of study that might be helpful, a pathological/epidemiological study. Such a study could take the following form:

(1) Assemble pathology slides from as many of the LLNL melanoma cases as possible, preferably beginning in 1970. For each case select the slide (if there is a choice) showing the most advanced lesion.

(2) Assemble twice as many slides, as in (1), from two other series of cases diagnosed during the same time period.

    (a) Melanoma cases among non-LLNL employees diagnosed in the Kaiser facility mainly serving LLNL population.

    (b) Melanoma cases diagnosed throughout the San Francisco Bay Area.

(3) Submit the slides from all three series in a random-blind fashion to three pathologists highly expert in melanoma with a request for classification as to advancement of the lesion, for example, Clark's levels and "thinness".

(4) From the reports of (3) ascertain the extent to which the excess in reported cases at LLNL can be accounted for by a tendency to diagnose relatively superficial lesions.

I believe that such a study would be highly useful in answering the first question."

"It is clear that during 1972-1980 the reported occurrence of melanoma at LLNL, but not at LANL, exceeded by three or four times the reported occurrence in the surrounding general population. It is also of interest that the incidence at LLNL in 1981 was considerably lower than in 1980 (the peak year), and that from September 1981 to September 1982 no cases were reported."

"At least four categories of possible explanations for the excess must be considered. There are differences between LLNL and the surrounding community with respect to:

1. Reporting of cases and enumerating the population from which they came

Conceivable discrepancies in the ways in which the cases (numerator of the rates) and the population (denominator) are counted could account for the excess. The quite complete reporting of cases and population in both LLNL and Alameda County, however, make this possibility unlikely to be a significant element.

2. The nature of the people

The LLNL employees have on the whole a higher level of education than the comparison population. Moreover, at some point during the period of excess, presumably about 1976, the LLNL group were highly sensitized to the occurrence of melanoma. These characteristics, and possibly others, might have led them to self-examination and to seek medical examination for melanoma more than was occurring in the general population.

3. The medical services provided

It is also possible that the general quality of services by the LLNL Medical Department and by the physicians in the community who served them (approximately half of the employees belong to the Kaiser Health Plan) was superior to that obtained by the comparison population. Again, about 1976 or earlier, sensitization of the physicians involved may have influenced them to search for and diagnose melanoma to a greater degree than their counterparts caring for the general population. For example, biopsy and pathologic examination of tissue may have been used more extensively. Legal considerations that evolved may have enhanced the tendency.

4. Factors in the workplace

From the outset of the studies workplace factors have been the obvious, paramount consideration. Although melanoma as an occupational disease must be a rare phenomenon, the matter still should be pursued. Thus far, it appears that working as a chemist or in certain buildings at LLNL possibly could be a factor in the occurrence of melanoma. Hence the issue merits further investigation with exploration of all significant leads."

Reviewer D:

"The initial report in Lancet 1981 by Austin, et al. (Lancet ii 712-716, 1981) compared the incidence of MM at LLNL to a local city and county incidence which may not be correct. The Department of Health Resources, headed by Dr. Austin, uses hospital based tumor boards and pathology department reports for accurate diagnosis of cancer. This method is presumably valid for internal malignancy where hospitalization and pathology review is inevitable. In the case of melanoma, however, the circumstances are different. Low risk melanoma can be and is diagnosed and treated by dermatologists without hospitalization and without the necessity of pathology department review. The numbers of such cases is unknown, but may be significant. The age adjusted rates for 1972-1977 (table 42) seem low considering the worldwide increase in incidence and mortality from malignant melanoma of the skin. Comparable cities by latitude in Australia have more than twice the currently recorded incidence. Dr. Austin's Department has never contacted the UCSF Melanoma Clinic for pathology reports to cross check with their own cases, ... Dr. Austin states that a previous check of Kaiser records, where the above under-reporting would not be applicable, showed no significant differences from the non-Kaiser incidence. This study has not been published and may not correctly address the problem."

"Several factors could contribute to the apparent increase reported by Dr. Austin. Study of the incidence of melanoma at the Los Alamos Facility (reference 63, 65) excluded individuals employed less than one year at the Los Alamos Laboratory. This seems to be a reasonable exclusion and has not been done in the Austin report. Secondly, the Hiatt and Fireman report of February, 1984, although confirming a higher incidence, made the observation that LLNL employees have significantly more skin biopses. This factor, together with the high education level, high skin cancer background, and high level of awareness of the entity of melanoma makes it extremely likely that early and precursor lesions would be discovered. This would correlate with

the four cases of dysplatic nevi initially diagnosed as melanoma and with the rather thin level of invasion of those tumors which were found after the initial 1972-1977 report. At the present time it is not certain whether the data available justify all the conclusions, but it seems that many of the recommendations based on the conclusions are valid to continue collecting data and following these patients."

"The apparent high increase in melanomas arising in moles raises the possibility that even the numbers of precursor moles may be elevated in the LLNL employees compared to background, a factor which may be difficult to control, as mole counts in normal populations are not currently available."

Reviewer E

"This is a comprehensive and substantial report of a thorough study of the possible factors underlying an apparent high rate of cutaneous malignant melanoma (CMM) in employees of the Lawrence Livermore National Laboratory (LLNL). I consider that the case for a real elevation in risk of CMM in employees of the LLNL is made by the data that I have seen."

Reviewer F

"Since this is the issue with which [the] Kaiser Permanente study dealt, it is of particular interest. [The Kaiser Permanente study] found that among Health Plan members, employees of the LLNL had more biopsies for skin lesions than their matched controls from Walnut Creek. Furthermore, there were more biopsies for pigmented nevi, especially of junctional nevi which are more similar to the most common form of malignant melanoma, the superficial spreading melanoma. There are two possible explanations of this finding: (1) that more malignant melanomas and pigmented nevi are being picked up (i.e., harvested) because of increased concern among employees and doctors, or (2) that there is no increased concern and there are simply more malignant melanomas and suspicious pigment nevi in the LLNL population that require biopsy. Either explanation is of interest. A third possible explanation is that there is both harvesting and a true increased incidence of melanoma. For the first to be true, either lesions which would eventually be diagnosed must be picked up sooner, or lesions which would never be diagnosed (i.e., they would regress or remain on a person until death) would have to be true. It seems most likely that if lesions would be eventually diagnosed and that they are simply being harvested early that the increase in incidence would be temporary. As their report points out, if harvesting per se were to have accounted for the excess, then (with an observed over 3

times that expected) new 'early diagnoses' during this interval should have accounted for more than the next dozen years of expected cases (p.45). Actually the next eight years would be accounted for but at any rate there is no evidence that the incidence of new cases has returned to pre-epidemic proportion despite 10 years of observation. Thus, the possibility seems unlikely that there has been harvesting of lesions which would eventually be diagnosed."

"The other possibility that some lesions are being picked up that would never be diagnosed is, I believe, a matter of debate and not 'rather remote'. There is no data on humans which says that malignant melanoma completely regress. This would be extremely difficult to ascertain in a clinical setting since one couldn't be sure an observed lesion has a melanoma until it was biopsed and if it was biopsed it would be removed or certainly changed by the biopsy procedure. However, there is some indirect evidence which can be brought to bear in favor of this argument. First, it is the natural history of pigmented nevi to regress. They appear in young adulthood, progress to dermal and some times pedunculated lesions and then disappear in later life (Stegmaier O. Natural regression of the melanocytic nevus. J Invest Derm 1959; 32:413-421). Also, spontaneous regression of primary melanoma is known to occur even though it is extremely difficult to document (Bodurtha AJ. In: Human Malignant Melanoma. Clark WH, et al., eds. New York: Grune & Stratton, 1979, pp.227-241). Second, there are well based theories of neoplastic development which hold that cells are transformed in one (or multiple) step(s) to a certain level from which they can go on to become frankly clinically malignant, they can regress back to normal or they can remain in that state until death. This first level is a state which could be interpreted as histological malignancy (Foulds L. Neoplastic Development. New York: Academic Press, 1969, pp.76-81). If lesions are biopsed at that stage, then it could be seen as an epidemic on a population-level. For this to be true in the present case one would expect lesions from LLNL to be thinner on the average than comparable to LLNL controls".

"It is very difficult to rule this possibility out because of our ignorance of the natural history of malignant melanoma. I do not believe it can be dismissed."

"I don't think a harvesting effect of "cancers" that may regress can be ruled out and these epidemiologic studies may be able to contribute much to this area where clinical and histologic data are sparse. I would predict that the current intensified surveillance program will result in a further increased incidence of malignant melanoma and that the incidence will remain high as long as the surveillance and awareness

persist. Will this be because of the surveillance or will a real increase of malignant melanoma due to occupational exposures be implicated? Two factors should hold if the increase in melanoma incidence in the past and the one now predicted are secondary to harvesting: (1) biopsed lesions should be thinner (i.e., earlier) than in a comparable population, and (2) case fatality rate for melanoma should be lower among LLNL employees than a control population."

Issue 2: If there is an excess of malignant melanoma at LLNL, can the excess be attributed to occupational exposures?

Reviewer A:

"The finding that chemists experience the highest of the occupational risks is of interest, particularly in light of previous observations. Like those observations, however, it is curious that whether one chooses personnel account information, general building affiliation, sources of skin exposure to chemicals, or use of protective measures to serve as alternative criteria of chemical exposure, it seems to be something about chemists other than their work with chemicals at LLNL that is involved."

"Both that positive finding, the finding of an effect related to exposure to sources of radiation, and that of an effect related to explosive fumes all appear to act as dichotomous variables without dose-response modification. Further, each of these seem to be common past exposures in this laboratory (over a third of the controls had a history of radiation exposure), and it is possible that any of these three factors may tell as much about a LLNL professional's career path as about his recent history."

"The crucial question for the laboratory is whether, after accounting for the personal and socioeconomic variables, there is any reason to suspect a specific laboratory activity or activities as a cause of the melanoma. I believe the most important table to answer this question is Table 39, which examines the residual significance of each after such an accounting. My interpretation of this table is that there still is substantial residual prediction for melanoma from a LLNL associated variable, and that the dichotomous effect of radioactive exposure is the best available, if unsatisfactory, choice. It would have been interesting to carry out this regression to one more step. Even though there would likely be insufficient power to adequately test whether an additional LLNL variable produced a risk incompatible with chance, the very size

of the risk estimate would suggest whether 'radioactive exposure', in combination with education and personal characteristics, could explain all the excess."

"In any case, since the evidence from outside investigations is very weak that radioactivity itself plays role in melanoma etiology, and since there is no evidence of a dose response, I find conclusion number 4 on page 58 to be too strong. From the data presented, one as yet unidentified occupational risk factor, when added to the precedented and plausible effects of 'constitution' and of education, plus chance, might provide an even better explanation of these 31 cases. That no such factor presents itself does not decrease one's dissatisfaction with those available."

"Most importantly, there is simply no unifying biological hypothesis which would explain these factors and still retain high plausibility in the larger context of melanoma etiology. If there is a cluster of melanoma at LLNL, which we must presume there is, a satisfactory explanation for it has not yet been identified."

Reviewer C:

"[Page 41 'A chance cluster']. The comment is made that distribution of MM risk is significantly aggregated within occupational groups. Individual occupations may have more than their share, but I found no analyses comparing the distribution of all cases and controls by occupation, and suggesting that the two were significantly different, either as raw data, or weighted for the previously known social distribution of melanoma risk. This is a somewhat important point, and the data should be available to the reader, rather than just the statement."

"[Executive summary page vi 'On the basis-'; page 116]. That the cumulative effect of 11 data items selected because of their power of discrimination between cases and controls in this data set should be able to account for 80% of the difference between them in the same data set does not seem surprising. Intuitively, it seems to me that if the 11 differences least likely to be due to chance were taken from a set of many comparison between two data sets that were in fact samples from the same large population, these 11 differences might well account for 80% of the variance. A formal discussion of the mathematical statistics would be helpful."

"Exposing people to ionizing radiation does not give them melanomas. There are almost no ways of producing melanomas in animals, in contrast to the large number for the other skin cancers."

"The robustness of the human melanocyte in resisting carcinogenic change provides the great scientific interest of the LLNL episode. However, it does militate against ready acceptance of the authors' findings. If, after years of essentially absent reporting of relationships between ionizing radiation or chemical agents and melanoma, in one small study three different ones (exposure to radio-active materials; the volative agent used in photography; and the agent associated with explosives) all appeared. With the tiny numbers, this multiplicity suggests a random distribution of the data, and throws the reader back to the statistical points raised above."

"The study has implicated exposure to fumes from photographic chemicals. This excess risk was only found among those occupationally exposed, and has not been reported previously in other circumstances. In the hurly-burly of medical practice, quite strong relationships can go unnoticed for decades. However, it is surprising that these occupational risks, if real, have not been reported previously."

"The excess risks among chemists and those with a higher degree are not unexpected, and by themselves the observations are too non-specific to indicate any policy beyond increased vigilance."

"The association with exposure to radioactive materials was not confirmed by the objective data of the badges. I do not believe that the exposure to radiation has been any greater at LLNL than at Los Alamos National Laboratory where a normal incidence of malignant melanoma has been found by a careful study using the same techniques as at LLNL. Given that there is indeed an increased incidence of malignant melanoma among the workers at LLNL, it seems highly likely that workers who had skin tumors would selectively recall exposures, compared with the healthy controls."

"In sum, the evidence about the 1969 building, work as a chemist, exposure to photographic chemicals or explosives, etc., suggest either complicated or generalized sources for the excess risk. In contrast, the limitation of the problem to the workers at LLNL argues for the importance of some precise chemical agent or industrial process not common elsewhere, and the present study has produced a number of broad and non-specific indications whose relative importance in uncertain. The ability of a case-control study to provide explanations as to why the cases became ill and the controls did not depends on the right questions being asked. The authors clearly made a major effort to do this. Either the problem at LLNL is indeed due to the independent operation of a

number of occupational factors, or the investigators were not lucky enough to ask about the right agent. Contemporary knowledge of environmental factors, apart from sunlight, in the causation of malignant melanoma is sketchy, and provides little guidance to the LLNL problem. Numerous case-control studies have been done by expert investigators in the past that yielded little information, notably on breast cancer and colon cancer. This kind of failure is simply a feature of the method."

"Thus those - the workers, the LLNL administration, the DOE, and those more generally involved with the melanoma problem - who are concerned are left by the continuation of the problem and the findings about its causation of the present study with little opportunity for action. They can do the vigilant things that the workers and Dr. Lawton are doing. They can in fact be more careful all along the line. But the findings do not have the necessary specificity or statistical power, either in the authors recommendations or in this reviewers opinion, for doing anything radical about the photography at LLNL, or the explosives handling, or Site 300. Following the authors' recommendations, we must wait until more cases have occurred, and then do a confirmatory case-control study."

Reviewer E:

"I have no reason to postulate that any of these constitutional characteristics would have been a likely confounder for the associations with occupational variables. It may be, however, that the positive association with parental history of skin cancer would have been explained by measurable constitutional characteristics had they been better measured."

"While I can think of no plausible basis for a protective effect of smoking against melanoma (nor have we observed one) I don't think that amalgamation of light smokers and non-smokers into one exposure category overcomes the 'problem'. As regards the occupational relationships, however, there would be no plausible basis for considering smoking as a possible confounding variable."

"These problems in measurement of exposure to the sun do raise the possibility of uncontrolled confounding in the occupational associations."

"The lack of increased risk of CMM in the technicians is anomalous. They might reasonably be expected to share the exposures of their laboratory principals and suffer the same or even greater risks."

"Understanding of the associations observed with particular job categories and this lack of association would have been aided greatly by an occupational hygiene assessment of the workplaces of particular categories of employees. Are the chemists, for example, more or less exposed to particular chemicals than the technicians that work with them? etc."

"Use of the 'building constructed in 1969' device for defining exposure is artificial and has little logical appeal. Again an occupational hygiene assessment would have been valuable. A more logical, exposure-based affinity between high risk work sites might have been identified."

"I don't find the 'point-source epidemic' postulate for risk associated with buildings 111 and 361 to be at all persuasive (see page 38). The variation in ORs on which this postulate is based (table 24) may well have been due to chance. It is interesting to note also that the same variation is seen for ORs for work in buildings built in 1966 although about a mean of near 1.0 rather than 3.1. The temporal variation therefore may lie in the reference category for these two exposures rather than in the exposed."

"There are occupational associations with CMM risk in these data for which no ready explanation is available at this time. Undoubtedly the strongest is with occupational exposure to radiation and there are other studies which support this association. [Comment by authors: the existence of these 'other studies' has not been documented]. The lack of support from the film-badge data, however, weakens the causal inference. Given the largely exploratory nature of the occupational aspects of the study, the problem of small numbers and doubtful validity of extensive modelling, probably all associations should be viewed as tentative and in need of further confirmation or refutation."

Reviewer F:

"Consistency. As the authors point out, consistency is one of the most compelling criteria by which epidemiologists infer causality. The data for consistency among persons exposed to ionizing radiation (which refers to the risk factors of being exposed to radioactive materials and to the Pacific Test Site) and then developing melanoma is almost nonexistent. On the other hand, as has been noted many times before, ionizing radiation does increase the risk of leukemia, breast and bone cancer which have shown no excess in the LLNL population. Also, the Los Alamos study showed no excess melanoma."

"Causality. I believe the contention that the identified occupational factors are 'causal' is stretching the sense of the word as it is usually used in epidemiology. The criteria that they list are accurate, but because the study was not testing any specific a priori hypothesis, because the role of chance is uncertain, and because consistency is weak, it seems too strong to conclude that the occupational factors are causal."

"My own feeling at this time is that there may well be some occupational factor(s) in the LLNL environment which have contributed to an excess of malignant melanoma, but that it is premature to call the items identified by this study 'causal'. The leads identified by this investigation need to be pursued in other settings and by other researchers. The recommendations take steps in that direction."

"Hundreds of p values are reported with no allowance for the notion of multiple testing. Then, in conclusions, authors state that they have shown that 'a variety of occupational risk factors are causally associated with MM among LLNL employees'. Worse yet, in the executive summary at the start of the report, they list 'five of seven appear to be independent contributors of risk' when several of them are shown in Table 39 to be non-significant. In fact, none of these factors are necessarily related to MM. If the effects of the non-occupational factors have not been adequately modeled, then the additional effect of the occupational factor may be spurious."

Issue 3: Were non-occupational factors adequately addressed?

Reviewer A:

"The findings with respect to personal and familial characteristics are properly discussed. I would agree that the cluster does not seem to be attributable to an excess prevalence of such factors at LLNL, although, again, it is impossible to rule out some contribution, given the number of observations."

"I find the leisure time/sun history results of great interest, and I have no alternative interpretations to those offered by the authors."

"While the health history component offers little positive explanation for the distribution of melanoma at LLNL, the imbalance of light smokers, and the suggestion of an increased past consumption of elective medical care among the cases do suggest some general, non-occupational behavioral differences between cases and controls. The finding of high risk associated

with a history of past personal and familial skin cancer is credible, though the finding is rather subject to recall bias and therefore necessarily should be taken with a grain of salt."

Reviewer C:

"[pps.82,83]. The excesses of parental non-melanoma skin cancer, and of earlier non-melanoma skin cancer in the cases, are interesting and attractive findings which appear to be new. Either nobody thought to ask, or the incidence of these cancers, which has been rising along with that of the melanomas, was previously too small for the relationship to emerge. However, the same problems of statistical significance in a situation of multiple testing emerge. They are probably less severe than in the occupational factors - the previous literature would lend strong support to a relationship between non-melanoma skin cancer and melanoma existing. Hence the current comparison is something of a test of an hypothesis. In contrast, the previous literature did not suggest a relationship between melanoma and ionizing radiation. Here the comparison advances a hypothesis de novo."

Reviewer E:

"Eye colour is well measured but the grouping of blue and green together is unwise. There is in my view no prior basis for believing that these two would indicate the same risk of CMM and their grouping may explain the anomalous results with eye colour."

"Hair colour appears to have been measured entirely by self assessment without the use of standard hair colour charts as a classification guide. This procedure would be error prone and likely to mask the expected positive associations with fair or red hair."

"Self-measurement of numbers of large naevi was probably also error prone and may have biased the resultant OR towards the null."

"I am doubtful whether acute skin response to sunlight after a few days previous sun exposure is a good measure of skin response to chronic sun exposure which is an important determinant of melanoma risk."

"Measurement of personal sun exposure is unsatisfactory. Specifically, limitation of recording of outdoor exposure periods to those accumulated after 21 years of age may have missed a critical exposure period (teenage and early adult life) in which recreational outdoor exposure is common and the future malignant potential of naevi may be determined."

"The 'protective' effect of total outdoor exposure in adult life is consistent with other data but may miss an important effect of intermittent exposure which is not addressed separately in the presentation of results in this study."

"The outdoor exposure measure also appears to have neglected outdoor work which, while not a feature of the LLNL work environment, may have applied during an earlier part of working life particularly in the non-scientist class."

Reviewer G:

"Number of large moles appears to be the most important predictor variable. It is not clear that splitting it into <6 and 6 or more moles adequately models the effect of moles on MM."

Issue 4: Were appropriate methods and analyses used?

Reviewer A:

"The protocol for conducting the case-control study seems quite appropriate. In particular, the choice of controls was appropriate in respect to both source and number. The use of both interview and record abstraction was wise."

Reviewer C:

"In these data, information on many items was available for each case and control. The analytic machinery made a very large number of comparisons. For some of these there was an a priori expectation of the way that the comparison would go. For example, it could reasonably be expected that more cases suffering from malignant melanoma than healthy controls would tan poorly and be susceptible to sunburn. That comparison is thus not part of a mass of repeated statistical tests, but it is already isolated by previous knowledge. The excess of cases with this characteristic was reported as significant at the 1 in 20 level, which is reasonable. However, for most of the items there was no previously available data to form a hypothesis that the cases would be different from the controls. Hence, for these items there is no way of knowing whether the statistical significance of a particular difference was an expression of the multitude of comparisons, or was real. When multiple comparisons are made, the individual probabilities are useful for identifying the differences most unlikely to be due to chance. Single ones cannot, however, be taken from their context, and reported as 'significant' findings. Techniques exist for correcting probabilities when multiple comparisons have been made.

Depending on the number of comparisons, they increase the estimated probability of any outcome, which naturally reduces the 'statistical significance'."

"The authors have chosen to present the statistical significance of various differences uncorrected for multiple comparisons. This exaggerates the certainty of their findings, so that they can discuss particular items as if their evidence supported solid conclusions. For example, in Table 27, 16 case/control comparisons are made for chemicals to which the skin of workers can be exposed, and in Table 28, 19 case/control comparisons are made for exposure to chemicals that can be inhaled. For two of these comparisons (the inhalation of volatile photographic chemicals and the inhalation of fumes from stored high explosives), the excess among the cases was statistically significant at the 0.05 level. In other words, for each of these two comparisons the probability that the difference could have occurred by chance was less than 1 in 20. However, 34 (16+19) comparisons were made. It is not at all surprising that in 35 comparisons, two differences will found of this degree of unlikeliness. Yet the authors devote considerable space to discussion of the biologic nature of the relationships, and of their interaction with other factors (pp.36-37). The photographic chemicals get a paragraph in the Executive Summary (p.v)."

"Adjustment for other variables from within the same data set (Tables 34 & 35, pp.104,105) does not address the multiple testing problem. Random variation would not be influenced by adjustment, which implies linkage between variables."

"The study under review confirms other information about the physical and behavioral characteristics of people who develop malignant melanoma, and succeeds in identifying factors that have been confirmed elsewhere, although they were only dimly perceived at the time it was designed. These welcome findings confirm the basic soundness of the conception and execution of the study. However, the principal purpose of the enterprise was to identify the cause or causes of the excess incidence of the disease in the working population at LLNL."

Reviewer E:

"This is a small study. There are only 31 cases. It is questionable, therefore, whether much weight can be placed on the multivariate analyses with their implied polychotomisation of the data and creation, probably, of many 'empty cells'. The small numbers may explain some comparatively large changes in odds ratio (OR) estimates from one model to another."

Reviewer F:

"The problem the study addressed was the need to search for some explanation of the previously described three to four fold excess in malignant melanoma at LLNL. No specific hypothesis was stated at the onset of the study and most epidemiologists would consider it an hypothesis generating study. This has some implications on the way findings can be interpreted and I'll mention this again later."

"Chance. Because this study was an hypothesis generating investigation, a multitude of comparisons between cases and controls was made. The authors point this out (p.ii) and state that 'significance values are presented without adjustment for multiple comparisons since the purpose of the investigation is an exhaustive analysis rather than an hypothesis test.' The problem with making multiple comparisons is that a certain number will be 'significant' just by chance, i.e., 5% of a set of twenty random comparisons will be different by chance and judged 'significant if one selects the 0.05 level of significance. It is difficult to know how many comparisons were actually made in this study, but to select out the several factors that were most strongly associated with melanoma by this process may have led the authors to place more importance on them than was justified in a strict statistical sense. I would agree that these observed differences have generated hypotheses about the occupational exposures to chemicals and radiation, but it seems premature to refer to them as causes. One cannot ever 'rule out chance', but can only say that a difference is likely to have occurred by chance with a certain probability."

Reviewer G:

"Despite the fact that the basic method of analysis is correct and appropriate, this report is less than impressive in terms of what was done and how the results were reported. Questions include:

1. Lack of a clearly expressed conceptual framework for performing the analyses. For example, in looking at exposures the last 10 years prior to the diagnosis was used implying a latent period for MM but this same period then was not applied to all other occupational data.

2. No discussion of the fit of the logistic model is ever given. Were goodness-of-fit tests done and what were the results?

3. When numerous variables are used, does the program give accurate estimates of the p values and of the odds ratio? The problem is the small sample properties of these estimates.

4. Numerous tests were performed that seem to be of little value, but then Table 39, which is critical to interpreting the outcome, is passed off in a brief paragraph. The remarks made in 3. above are of concern here."

"Is it true that both the cases and controls could equally remember events occurring years ago and immediately prior to the diagnosis of MM for the case? Would the employee records be equally complete for cases and controls? No discussion of checks on the validity and reliability of the data are given."

"No clear discussion of how confidence limits were obtained. Were exact methods used? Why not give asymptotic estimates or correlations instead of so many odds ratios?"

"If you try to discriminate the two groups with one of the logistic regression equations, how do you do? Can a high proportion of the cases and controls be correctly classified?"

APPENDIX B
LIST OF REVIEWERS

1. Bruce Armstrong, Director
   NH & MRC Research Unit in Epidemiology and Preventive Medicine
   University Department of Medicine
   The Queen Elizabeth II Medical Centre
   Nedlands, Western Australia 6009

2. Lester Breslow, M.D., M.P.H.
   Professor of Public Health and Director of Cancer Control Research
   Division of Cancer Control
   Jonsson Comprehensive Cancer Center
   10920 Wilshire Blvd., Suite 1106
   Los Angeles, California 90024

3. Virginia A. Clark, Ph.D., Professor
   Department of Biostatistics and Biomathematics
   School of Public Health 51-253-CHS
   University of California, Los Angeles
   Los Angeles, California 90024

4. Robert A. Hiatt, M.D., Ph.D.
   The Permanente Medical Group, Inc.
   Department of Medical Methods Research
   3451 Piedmont Avenue
   Oakland, California 94611

5. J.A.H. Lee, M.D., Professor
   Department of Epidemiology
   School of Public Health and Community Medicine
   University of Washington
   Seattle, Washington 98195

6. Thomas M. Mack, M.D., Professor
   Department of Preventive Medicine
   School of Medicine
   University of Southern California
   2025 Zonal Avenue PMS 5105
   Los Angeles, California 90033

7. Richard W. Sagebiel, M.D.
   Professor of Pathology and Dermatology
   Co-Director, Melanoma Clinic
   University of California Medical Center
   San Francisco, California 94143

REFERENCES

1. Acquavella JF, Tietjen GL, Wilkinson GS, Key CR, Voelz GL. Malignant melanoma at the Los Alamos National Laboratory. Lancet 1982; i:883-884.

2. Austin DF. A study of cancer incidence in Lawrence Livermore National Laboratory Employees. Report #1. Malignant melanoma. Department of Health Services, Resource for Cancer Epidemiology Section, Berkeley, California: April 17, 1980.

3. Austin DF. A study of cancer incidence in Lawerence Livermore National Laboratory Employees. Report #2. All sites of cancer. Department of Health Services, Resource for Cancer Epidemiology Section, Berkeley, California: February 4, 1983.

4. Austin DF, Reynolds P. A case-control study of malignant melanoma among Lawrence Livermore National Laboratory employees. Unpublished. Department of Health Services, Resource for Cancer Epidemiology Section, Berkeley, California: July 3, 1984.

5. Hiatt RA, Fireman B. Malignant melanoma at Lawrence Livermore National Laboratory: incidence in a prepaid health plan and a case-control study. Unpublished. Presented in part at the annual meeting of the Society for Epidemiologic Research, Houston, Texas, June 13-15, 1984.

6. Moore D, Bennett D, Mendelsohn M. Melanoma among LLNL employees: an epidemiologic puzzle. Paper prepared for submittal to Statistics Symposium on National Energy Issues, Seattle, Washington, October 16-18, 1984, UCRL-91475 Preprint. Livermore, California: Lawrence Livermore National Laboratory, 1984.