Attachment 5.7

The Epidemiological Evidence on the Carcinogenicity of Beryllium in Humans

Brian MacMahon, MD, PhD

Authoritative reviews of the question of whether occupational exposure to beryllium compounds is associated with increased risk of respiratory cancer were published in 1987 and were critical of the quality of the evidence available up to that time. No clear conclusion could be drawn from it as to the carcinogenicity of beryllium to humans. If studies published since 1987 are to lead to a revision of the regulatory status of beryllium compounds they must clearly be of high quality and scientific validity. These studies, as well as the earlier reports, are reviewed here. I argue that the small and inconsistent excess of lung cancer deaths in employees of one or two plants seen in two post-1987 studies is compatible with a number of explanations other than that they are attributable to occupational exposure to beryllium. Specifically, information on cigarette smoking is poor, and the data do not exist to rule out the possibility that the small number of excess deaths results from residual confounding by cigarette smoking patterns in the populations studied. Indeed, excess deaths from emphysema and ischemia heart disease in the same cohort suggest that confounding by cigarette smoking is a more likely explanation of the lung cancer excess than is occupational exposure to beryllium compounds.

From the Department of Epidemiology, Harvard School of Public Health, where the author is Henry Pickering Walcott Professor of Epidemiology, *Emeritus*.

Address correspondence to: Brian MacMahon, M.D., PhD, 677 Huntington Avenue, Boston, MA 02115

0096-1736/94/3601-0015$03.00/0

Since the mid-1930s the relationship between exposure to beryllium compounds and various nonmalignant diseases of the lungs—particularly acute chemical pneumonitis and chronic berylliosis—has been the object of intense investigation of its clinical, epidemiological, and immunological aspects.[1] Considerable progress has been made, leading to an important understanding of the pathogenesis and practical methods for the prevention and treatment of this group of diseases.[2] On the other hand, fewer investigators have concerned themselves with the possible link between beryllium exposure and malignant disease of the lung, and progress in understanding this relationship, if indeed it exists in humans, has been nonexistent.

There is a useful dichotomy with respect to the epidemiological studies of cancer and beryllium exposure in humans between those reported before 1987 and those reported subsequent to that year. Studies of beryllium exposure and lung cancer risk undertaken between the mid-1960s and the early 1980s were reviewed by a number of national[3] and international[4,5] groups in 1987 and were found to provide an inadequate basis for coming to a judgment as to the carcinogenicity of beryllium for humans. Two more recent studies, both under the auspices of the National Institute for Occupational Safety and Health (NIOSH),[6,7] have been of good quality, but the interpretation of their findings is controversial. The International Agency for Research on Cancer (IARC) is reviewing once again the question of the carcinogenicity of beryllium compounds in

both humans and other species. The current IARC classification of beryllium and its compounds with respect to carcinogenic risk to humans (without further specification of the compounds) is 2A ("probably carcinogenic to humans"), a judgment based primarily on the accepted carcinogenicity of some beryllium compounds in some species of laboratory animals by some routes of administration. Given the rather unfavorable reviews of the epidemiological work appearing before 1987, it seems that any change in this classification must be highly dependent on the studies published since that time, and this review will focus on the epidemiological studies published since 1987. However, the draft IARC report suggests that the post-1987 studies represent only the latest contribution to an accumulating body of evidence on the topic. Before focusing on the latest studies, it seems necessary therefore to consider the evidence available prior to 1987.

Studies Prior to 1987

The Studies of Mancuso and His Colleagues

The early studies of Mancuso and his colleagues[8-11] must be mentioned, although a poor description of methods and inadequate statistical technique make almost valueless the information to be derived from these studies. Mancuso deserves credit for drawing attention to a resource that only in his time was becoming available. He used the records of the Social Security Administration (SSA), established in 1937, to create lists of workers employed in two beryllium companies in Ohio and Pennsylvania from the Quarterly Earning Reports (QERs) filed under law by the companies. The plants were the Lorain, OH, plant of the Brush Beryllium Company (now Brush Wellman, Inc.) and the Reading, PA, plant of Kawecki Berylco Industries (now NGK Metals of Japan). Through records of continued payments or filings for death benefits, reasonably complete ascertainment could be made of the vital status of the persons on these lists. Enquiries of the states wherein the deaths occurred led in a high proportion of cases to the obtainment of copies of death certificates, including classifications of causes of death.

In the first of these studies, 3685 white male employees who were reported on QERs between 1937 and 1948 were assembled into successive 3-year cohorts according to year of first employment and followed for 22 years or to 1966.[8] The number of deaths overall and cause of death were compared between the two plants and with those in a cohort of workers in the rubber industry, also assembled from SSA files. The high turnover rate of the beryllium workers was noted—only 22% worked for more than 2 years. This paper lays stress on the difficulties and inaccuracies of these sources of information (when used as sole source documents) and the frequent discordance between information from death certificates and that in the Beryllium Case Registry, established in 1952. The paper also reveals the authors' lack of experience with appropriate methods of statistical analysis of this kind of material. With respect to lung cancer, the only observation of note is of a "slightly higher" rate in each of the beryllium plants in comparison to that in the rubber workers. No significance tests were applied, and neither in this nor in any other of Mancuso's studies was any account taken of the possible confounding effect of smoking. Nor in any of Mancuso's papers was any estimate made of the completeness of mortality ascertainment through the SSA system.

Further analyses of these data published in 1970[9] and 1979[10] are practically uninterpretable because of the methodological problems referred to above. However, it did appear that the risk of lung cancer was higher among those who worked in 1–5 quarters than in those who worked for 6 or more quarters—i.e., in short-term employees. Moreover, within the category of employees reported in only 1–5 quarters, the risk seems to be concentrated among those working the fewest number of quarters—a phenomenon that seems to be directly opposite to that expected if it is to be explained in terms of occupational exposure to beryllium. Mancuso also claimed, based on 6 deaths among 142 cases of respiratory illness reported to the Atomic Energy Commission, that patients with benign beryllium diseases were at particularly high risk for lung cancer.

The follow-up of the two beryllium cohorts was updated to 1976 and reported in 1980.[11] A cohort of employees in the viscose rayon plant over the same period was used as a comparison group. Mortality rates specific for age and duration of employment derived from the viscose rayon cohort were applied to the beryllium cohorts to estimate the expected number of deaths from lung cancer. Two sets of expected numbers are shown—one based on all members of the rayon cohort and the other on members of the cohort who stayed in the same department throughout their employment in the plant. Use of these two expected values brings an unnecessary complication into the analysis, and they do not resolve the problem it seems they were intended to—that is, that the high lung cancer risk of beryllium workers may be related to some characteristic or characteristics associated with short-term employment. However, it is short duration of employment in the plants that characterizes the beryllium workers, not short durations of employment in one department of a plant. Nevertheless, the number of lung cancer deaths in the beryllium cohorts exceeds both expected values by about 50%. The differences between expected and observed are said to be statistically significant, but the significance is overestimated by the method used, since no account is taken of the variability in the rates in the rayon cohort that are used to compute the expected values. Reproducing tables from another, not readily accessible publication in which U.S. rates were used to compute expected values,[10] the author shows that the excess of lung cancer deaths occurred only after latencies of 15 years or more, although it should be noted that this experience accounted for the great majority of the total person-years.

The Mancuso studies, and in particular the 1979 and 1980 studies, are

criticized in detail in the Environmental Protection Agency (EPA) 1987 *Health Assessment Document for Beryllium*.[3] The document suggests that at least some of Mancuso's analyses used the NIOSH life-table program, although there is no mention of this in the Mancuso papers. If the NIOSH program was used, then the expected numbers are underestimated by approximately 11% because of use of inappropriate population death rates, as discussed below in the context of the study of Wagoner et al.[12] Dr. Jean French, a NIOSH epidemiologist, reported that NIOSH reanalyzed the Mancuso data on the beryllium and viscose rayon workers and found "serious problems" with the Mancuso analysis. The EPA was unable to obtain the results of the analysis from NIOSH, and the discrepancy could not be resolved.[3 pp 7-50] The EPA's evaluation also expressed reservations about the adequacy of the age adjustment in Mancuso's analysis but because of lack of access to the data was unable to determine the magnitude of the effect of the deficiencies. The report also comments on the lack of discussion of the confounding effects of smoking. EPA concludes:

> . . . despite the author's certainty regarding the existence of a causal relationship between beryllium exposure and lung cancer, the evidence presented in this study is not convincing because of the many limitations of the study, as described above. Hence it would appear that the study is at best only suggestive of an increased risk of lung cancer due to exposure to beryllium. (pp. 7-51)

The NIOSH Studies of Occupational Cohorts

A series of studies of workers employed in beryllium production facilities was first reported by Bayliss et al. in 1971.[13] Initially, the series encompassed 10,356 current and past employees of the Brush Beryllium Company and Kawecki Berylco Industries, the only North American producers of beryllium. After removal of 2153 workers whose records were inadequate, 1130 females and some other smaller groups, a study cohort of 6818 males was assembled. The cohort was followed through 31 December 1967, and 777 deaths were identified. This number of deaths was fewer than the approximately 842 expected on the basis of death rates for males in the U.S. population, a fact that was attributed to the healthy worker effect. There was no elevated risk of lung cancer overall in the cohort (obs = 36, exp = 34.1) and no excess risk in relation to length of employment, date of first employment, or whether the employee worked in an office or in production. However, the data were not analyzed for a relationship to latency (the interval between first employment and the period of observation), a variable on which subsequent investigators place considerable emphasis. There were no significant excess risks from any other cancers in this cohort.

Subsequent versions of this study[12,14,15] excluded more and more plants, until the latest (1980) version was based solely on the employees of a single large plant in Reading, PA.[12] This paring down was said to be done for the purposes of improving the quality of the data and to ensure sufficient numbers of employees with long latent periods. These are valid reasons for such an operation, but it is difficult to ensure—and to convince others—that in the process, when the contents of the data are known to the individual doing the paring, that sets of data that were supportive of a particular hypothesis were not preferentially retained over sets that did not.

The latest version of this study[12] is based on 3055 white males employed for at least 1 day between 1 January 1942 and 31 December 1967. Follow-up was attempted through a variety of mechanisms to 1 January 1976 and was completed for all but 3% of individuals. Eight hundred seventy-five deaths were identified (817 expected), of which 47 were attributed to lung cancer (34.3 expected). The excess of deaths from lung cancer was statistically significant. Observations with a latency period of 25 years or more showed an almost doubling of the lung cancer risk, which was statistically significant. A latency of 15–24 years was associated with an excess number of cases (obs = 18, exp = 14.1), which was not, however, statistically significant. The excess was limited to workers employed for less than 5 years, but the numbers of men employed for longer periods was quite small. There was no excess of cancers other than of the lung—indeed, the total numbers of cancers (including lung) were quite similar (obs = 143, exp = 136.2).

As noted, this paper was finally published in 1980, but successive drafts began to appear in 1977 and formed the crux of the epidemiological evidence before the OSHA Administrative Law Hearing in 1977.[15] The drafts and the final paper were extensively criticized by industry consultants, by an Internal Review Group within the Centers for Disease Control (NIOSH's parent agency), and by the original senior author himself.[16] These and other criticisms are assembled and extensively addressed in the 1987 EPA Health Assessment Document. The principal points are as follows:

1. Annual lung cancer death rates in the general population were not available at the time of the study for the years after 1967. For the period 1968–75, the authors assumed that the rates for 1967 prevailed. Since lung cancer rates were increasing during this period of time, this resulted in an underestimate of the expected number of lung cancer deaths, the size of which the EPA (1987) estimated at 11%, a figure close to that estimated by consultants to Brush Wellman, using an independent computer program. Correction for this error alone was sufficient to eliminate the statistical significance of the difference between expected and observed numbers of deaths.

2. The effect of confounding by the higher prevalence of cigarette smoking in the cohort relative to the general U.S. population was underestimated. A Public Health Service survey of 379 employees currently employed at the plant in 1968 indicated that the frequency of heavy smokers among them was higher than in the national data (21.4 and 15.3%, respectively). Wagoner et al. estimated that the distri-

bution of cigarette smoking patterns in the plant, compared to the U.S. general population, might have increased the risk of lung cancer by 14%, which they dismissed as an explanation of the increased lung cancer risk. The estimate of 14% is highly dependent on a number of assumptions, including those of the relative risks associated with various smoking categories. Wagoner et al. used the data of Hammond[17] for this purpose, but there are many other possibilities. Depending on the assumptions, the EPA (1987) estimates a range of 4.6 to 18.8% in the underestimate of expected deaths.

3. One individual who died of lung cancer was added to the cohort in a late draft of the paper, probably on the basis of his appearance on a QER. Company records show that he was hired for the purpose of taking a preemployment physical examination, which he did not pass. He was terminated the same day but was paid for the time spent in the examination and therefore was reported on the QER. His company record (of which NIOSH had a copy) clearly states that he "did not pass chest x-ray." Bayliss originally excluded this individual from the cohort, but the individual was subsequently added back to the cohort by the other investigators. The addition of this one case makes the difference between observed lung cancer deaths and the authors' own estimate of expected deaths formally significant, whereas without it the difference is not formally significant. EPA presents a table in which this individual has been excluded, the expected values adjusted for the underestimation of general population rates and for the 4.1% (the minimal estimate) due to cigarette smoking patterns. Overall, the expected number of lung cancers is 41.9 and the observed is 46. There is no significant difference between observed and expected values in this table for any category of latency.[3(pp. 7-36)]

There are too many other criticisms of this paper (amply reviewed in the EPA document) to recount here. Perhaps most disturbing, however, is the impression one gets from the successive drafts of the paper of the authors' determination to "make a case" for the carcinogenicity of beryllium, rather than to objectively assemble the facts. For example, the nonemployee just referred to remains in the final publication, although the inappropriateness of this was pointed out to the investigators at an early stage. He was the only individual added between two successive iterations of the data set. The underestimate of expected deaths because of use of inappropriate population mortality rates was not acknowledged by the authors at first, even after it had been pointed out to them. This tendency of the authors was noted not only by outside reviewers of the successive drafts but by the CDC Internal Review Committee in several memoranda to the authors. One such memorandum, for example, refers to

> . . . the authors' tendency to under-emphasize, attack, or explain away, any findings which are inconsistent with the hypothesis that beryllium is carcinogenic. (p. 2)

and

> After repeated requests by the Panel, there has been some "toning down" of the paper's commentary, but the enthusiasm of the authors for the position that beryllium is unequivocally carcinogenic in man is readily apparent to the reader. (p. 3)[18]

The EPA document reads as follows:

> To summarize, it appears that the authors of the Wagoner et al. (1980) study tended to exaggerate the risk of lung cancer in the population of workers potentially exposed to beryllium, and underemphasized or did not discuss sufficiently the shortcomings of the study. The net effect was to turn a "nonsignificant association" of lung cancer with beryllium exposure into a questionable "significant association." (pp. 7-39)

The Beryllium Case Registry Study

Infante, Wagoner, and Sprince undertook in 1980 a follow-up study of patients enrolled in the Beryllium Case Registry.[19] This study is mentioned here for the sake of completeness, since it was part of the evidence considered by review groups during the 1980s. However, numbers were small, and the evidence in isolation would have carried little weight. The study has been extended by Steenland and Ward,[6] and a description of the up-dated study appears below.

Summaries of the Evidence Prior to 1987

Both the EPA and IARC reviewed in 1987 the evidence with respect to the carcinogenicity of beryllium in humans. It is clear that neither agency was willing at that time to categorize beryllium as carcinogenic to humans. The IARC *Monographs*, both in 1980[4] and again in 1987,[5] categorized "beryllium and beryllium compounds" into category 2A ("probably carcinogenic to humans"), primarily on the basis of the animal studies. The evidence of carcinogenicity of beryllium in humans was described as "limited." The EPA in 1987 used a somewhat different terminology. To quote from the EPA's summary[3]:

> Epidemiologic studies provide equivocal conclusions on the carcinogenicity of beryllium and beryllium compounds. Early epidemiologic studies of beryllium exposed workers (see IARC, 1972, 1980; Bayliss et al., 1971; Bayliss and Lainhart, 1972) do not report positive evidence for increased cancer incidence. However, recent studies do report a significantly increased risk of lung cancer in exposed workers. The absence of beryllium exposure levels and a demonstrated concern about possible confounding factors within the workplace make the reported positive correlations between beryllium exposure and increased risk of cancer difficult to substantiate. This relegates the reported statistically significant increase of lung cancer to, at best, an elevated incidence that is not statistically significant. Because of these limitations, the EPA (U.S. EPA, 1986) considers the available epidemiologic evidence to be "inadequate" to support or refute the existence of a

carcinogenic hazard for humans exposed to beryllium.

This designation of the epidemiologic data as "inadequate" differs from that of the International Agency for Research on Cancer (IARC, 1980) which concluded that the epidemiologic data provides "limited" evidence for the carcinogenicity of beryllium. In the EPA evaluation, more recent unpublished tabulations and analysis of the earlier study cohorts that correct for errors in the data base and the National Institute for Occupational Safety and Health (NIOSH) Life-Table program were included. Use of this newer data provides a basis to change (in a direction away from a judgment of carcinogenicity) the weight-of-evidence conclusion for the human data. (pp. 2-7) (parentheses added)

It should be pointed out that the "earlier studies" referred to in the second paragraph of this quotation are not the "Early epidemiologic studies" referred to at the beginning of the first paragraph, but rather the study of Wagoner et al. published in 1980,[12] which was the mainstay of such evidence as there was at that time for the carcinogenicity of beryllium in humans. The "errors in the data base and the ... NIOSH ... Life-Table program" refer specifically to that study.

My own evaluation of the evidence up to 1987 is perhaps evident from the comments above on those specific studies that had presented positive associations. The EPA's term, "inadequate," seems a better descriptor of the material than the IARC's term, "limited." It follows that studies published since 1987 must carry considerable weight if they are to be the basis for a reclassification of beryllium as a human carcinogen.

Studies after 1987

There are only two important epidemiologic studies published since 1987 relating to lung cancer risk in beryllium-exposed persons.

Steenland and Ward, 1991

Steenland and Ward followed patients enrolled in the national Beryllium Case Registry (BCR) since it was opened in 1952.[6] In the late 1970s, when the registry was copied by NIOSH, it contained 888 records. From these, and a "handful" added to the registry after 1980, patients who were dead at the time of enrollment or missing crucial follow-up information, and a few whose disease or exposure status were uncertain, were excluded, leaving 689 who were followed. It is unclear when accessions to the registry ceased, but 70% of the patients were enrolled before 1960 and 89% before 1970. Follow-up was by a number of national record-linkage resources and attempts at personal contact. Ninety-five percent were followed to death or were known to be living in 1988. Numbers of deaths by cause were compared with expected values based on U.S. rates specific for age, race, sex, and calendar time. In most of the earlier studies described above the estimate of the ratio of the risk in the exposed cohort to that in the unexposed or comparison group has been referred to as the "relative risk" (RR) or "risk ratio" (RR). In the more recent studies now to be described, the authors have used the term "standardized mortality ratio" (SMR) to identify the same ratio, and we have retained their terminology. In this context, RR and SMR can be used interchangeably.

Not surprisingly, there was a very large number of deaths from pneumoconioses and other respiratory disease; these accounted for 37% of all deaths—34 times the expected number. There were 28 deaths from lung cancer, compared to 14 expected, giving an SMR of 2.0 with CI = 1.3-2.9.

There was no apparent association with latency in these data. The SMRs for person-years within 20 years of exposure and those for person-years more than 20 years after exposure were similar: 2.0 (CI = 0.9-3.6) and 2.0 (CI = 1.2-3.2), respectively. This finding is at variance with the findings in other studies in which positive associations have been found; it is not commented on by the authors.

Data on tobacco smoking were obtained as of 1965 for 32% of the cohort. They suggested that the cohort had a lower prevalence of current smokers (for males, 41% compared to 54%) but a high percentage of former smokers (for males, 33% compared to 22%) compared to age-adjusted frequencies in a sample of the U.S. population. As the authors indicate, the differences are difficult to interpret because the presence of respiratory disease in the BCR patients may have discouraged smoking.

The authors suggest that patients registered with acute beryllium disease had higher SMRs for lung cancer (n = 17, SMR = 2.3, CI = 1.4-3.7) than those registered with chronic disease (n = 10, SMR = 1.6, CI = 0.8-2.9). Indeed, the SMR for patients with chronic disease does not significantly exceed 1.0. However, the numbers are small and the confidence intervals of the two estimates overlap. Eisenbud[20] has pointed out that the experience of the acute cases with respect to lung cancer, at least as reflected in the BCR, could be related almost entirely to a single plant, the Lorain plant of the Brush Beryllium Company. Eisenbud checked the employment records of the 17 lung cancer deaths that occurred in acute cases and found that all were reported from Brush plants and all but one from the plant at Lorain, OH. The latter case was reported from another plant but had also worked previously, in fact for a much longer period, at Lorain, so that all 17 cases had experience at Lorain. Eisenbud has records of 235 total patients registered with acute disease (Steenland and Ward give 237), of which 145 came from Brush; there remained 90 patients registered with acute disease from sources other than Brush, none of whom died of lung cancer.

Steenland and Ward, in response, report yet another set of values from this experience which differs slightly from both the numbers given by themselves previously and from those given by Eisenbud.[21] They offer an explanation as to why their latest numbers differ from those reported from Lorain in the cohort study of Ward et al., but more relevant is the question of why these numbers differ from those reported by Eisenbud. Apart from the absence of plant identification for some cases in the data

set of Steenland and Ward, there may also be differences in whether the two groups of authors are assigning cases to plants on different bases (e.g., plant from which data were reported, or something else). In any event, whether one takes for deaths from lung cancer among acute cases the 0/90 of Eisenbud or the 2/78 (RR = 1.0, CI = 0.1–3.6) of Steenland and Ward, it is evident that there is no higher lung cancer risk for the cases of acute (as opposed to chronic) disease other than in subjects from the Lorain plant.

After some discussion of the possibility of excess smoking by this cohort, and the possibility of selection bias, the authors conclude that "A more probable explanation for the observed lung cancer excess in this cohort is that these patients had received high doses of beryllium, a suspected lung carcinogen." The authors favor this explanation because of the overall excess of lung cancer in the cohort, relative to the U.S. population, and the higher risk in the patients with acute disease who are presumed, with justification, to have had the higher exposures.

However, this study seems a slim reed on which to rest a case for beryllium carcinogenicity in humans. The alternative explanations for the excess—smoking and selection bias—are dealt with by the authors quite superficially. Thus, with respect to smoking, which must be considered a critical confounder when dealing with relative risks as low as those found here:

1. Reliance is placed on data for a single year, 1965, a year chosen "because U.S. survey data were available at the time and because smoking habits in the 1960s are considered to have been most relevant for lung cancer mortality in the 1980s." Why the 1980s were thought to be the most critical years is unclear. The average year of first exposure of the members of the cohort was 1944. Neither the distribution of person-years of observation over time, nor the average year of death is given, but with exposures beginning in the 1940s and a presumed latency of 15 years one would suppose that any period from 1960 on would be relevant. One cannot in fact characterize the relevant smoking experience of this cohort on the basis of data from any single year, considering the major trends that have taken place in smoking prevalence in the population and the changing size of the group under study over this time period.

2. The Interview Surveys carried out by the National Center for Health Statistics, from which the U.S. population data were derived, use in-person interviews and standardized questionnaires. Smoking information for the registry patients came from "direct interviews or interviews with next-of-kin or from registry records." The proportions from each source are not given. It is not appropriate to compare histories deriving from such diverse sources.

3. There are variations in smoking prevalence in the United States by geographic region and by social class, of which no account was taken.

4. Smoking information was obtained for only 32% of the registrants, and no attempt to evaluate whether this small sample was representative of the whole is reported.

5. Finally, and most devastatingly, the smoking history of the registrants related to the time when they already had respiratory disease and had been reported to the BCR. Although this is mentioned by the authors, the total inappropriateness of comparing smoking prevalence in such a group with that of a sample of a generally fit population is not acknowledged. The period of life during which the smoking habits of this cohort were most relevant to their ultimate risk of lung cancer is quite certainly the period before they developed benign beryllium disease, not after.

The issue of selection bias is dealt with by the authors by noting that

1. Individuals who had died before entering the registry were excluded.

2. Only five individuals were known to have cancer when they entered the registry, and none of these had lung cancer.

3. If patients with lung cancer had entered the registry preferentially, one might expect their follow-up period to be short. However, only three of the 28 lung cancer deaths occurred within 5 years of entering the registry, and the lung cancer patients had an average length of follow-up (21 years) similar to that of the total cohort (20 years). This last observation, of course, raises its own questions, such as why do the lung cancer cases not have longer follow-up if risk increases with latency, as it is said to do in other studies?

In fact, no one really knows much about the selective factors that led to enrollment in the BCR, which was voluntary. It is known that there was considerable variation between plants and other sources of cases in (a) the diagnostic criteria for beryllium disease and (b) the probability of a case being registered if diagnosed. Steenland and Ward raise the possibility that "decedents with lung cancer may have been preferentially referred to the registry."[5(p 1381)] If that was so, there seems no reason why it should not also have been true for lung cancer cases registered while alive, although the few deaths in the first 5 years do argue against it.

The authors' belief that the risk of lung cancer death was higher for the patients registered with acute disease than for those registered with the chronic form is not well supported. The numbers are small and the difference between the two SMRs is not statistically significant. Furthermore, its dependence on acute cases reported from Lorain, as noted by Eisenbud, raises questions that I will return to below.

Ware, Okun, Ruder, Fingerhut, and Steenland, 1992

Ward et al. report on the mortality of 9225 males employed for at least 2 days between 1940 and 1969, inclusive, at any of seven beryllium production or processing plants in the United States.[6] The cohort was followed through 31 December 1988 by matching against a number of national databases. Vital status was known as of the study end date for 8921 (97%) of the cohort. Life table analyses compared observed deaths by cause with expected numbers computed from the U.S. general population rates specific for race, sex, age, and time period. The influence of geo-

graphic variation was evaluated by using county rates to generate other expected numbers for the period 1950–88, county data not being available for the earlier years. To evaluate the possible effect of unusual patterns of cigarette smoking in the cohort, information was obtained on members of the study population who participated in a medical survey conducted by the Public Health Service at four of the participating plants in 1968. The members of the cohort who participated in the USPHS survey were compared with the samples of the total U.S. population from two national surveys undertaken in 1965 and 1970. To determine how much the SMRs in the whole cohort would be changed if the cohort had had the same distribution by smoking habits as did the U.S. population, the authors use a formula for computing the overall SMRs for nonsmokers and light, heavy, and former smokers derived from an external source weighted according to the proportions of the study population in each smoking category. A "correction factor" was derived to estimate how much higher (or lower) the rates in the cohort would have been if their distribution by smoking patterns had differed from those in the general population.

In the cohort overall, 280 deaths from cancer of the trachea, bronchus, and lung ("respiratory cancer") were identified, compared to 222 expected, giving an SMR of 1.26 and a CI of 1.12–1.42. Data for two small plants in the same city were combined, making six "plant groups," plus one group consisting of men who worked in several plants and one comprising those for whom the specific plant was unknown. Among these eight groups, respiratory cancer deaths were significantly elevated only in the two oldest plants: Lorain, OH (obs = 57, exp = 33.7, SMR = 1.69) and Reading, PA (obs = 120, exp = 96.8; SMR = 1.24).

There was no trend with duration of employment. Indeed, the only excess of observed over expected that was statistically significant was among workers employed for less than 1 year. This was, however, the largest category of employee.

There were increases in SMR with increasing latency (interval since first employment). No excess of lung cancer deaths was observed within 15 years of employment. After a latency of 30 years or more, the SMR in the cohort overall was 1.46; at Lorain it was 1.66, and at Reading 1.40.

The SMR was highest (1.42) for workers first employed before 1950, but, as noted below, interpretation of this observation is complicated by the changing distribution of the work force over time: for example, Lorain, the plant with the greatest lung cancer excess, burned and was closed in 1948 and had no workers first employed after 1949. For workers employed in 1960 or later, in the cohort overall the risk of death from lung cancer was actually significantly below expectation (SMR = 0.62, CI = 0.36–0.95).

Application of the "correction factor" described above to adjust for smoking behavior in the cohort suggested that cohort rates should be divided by a factor of 1.1323 to accommodate the more frequent smoking prevalence in the cohort compared to the U.S. population. When applied to the SMR for the cohort overall, the SMR was reduced from 1.26 to 1.13; that is, the *excess* relative risk of lung cancer (0.26) was halved. The authors applied the same factor to compute adjusted SMRs for Reading and Lorain, but we should note that Lorain, which was closed at the time of the survey, was not even represented in the Public Health Service sample. With the "correction" the SMR for Lorain remained significantly elevated, but that for Reading did not.

This report must be considered against the background of the 1980 study of the plant at Reading. In the new study many of the deficiencies of the earlier work have been corrected: the study cohort has been carefully defined, the use of a 2-day minimum period of employment eliminates the problem of inclusion of individuals "hired" only for the purpose of taking pre-employment physical examinations, and follow-up and acquisition of death certificates were unusually complete, even for the plant where follow-up was least complete (Lorain). The life table used to compute expected values uses appropriate general population rates, and at least an attempt has been made to deal with the very difficult problem of geographic variation in lung cancer rates—difficult because workers frequently do not live in the same county as that in which the plant at which they are employed is located. Altogether, I have the impression that the data for this report were scrupulously collected and expertly analyzed.

It is in the interpretation of the findings that I find reason to differ with the authors, and in particular with respect to the statement in the authors' abstract that "Occupational exposure to beryllium compounds is the most plausible explanation for the increased risk of lung cancer in this study" (p. 885). Not only do I believe that a pattern of heavy cigarette smoking is a more likely explanation of the lung cancer excess, but the data seem to exhibit a number of characteristics that appear to be in conflict with the idea that beryllium carcinogenesis is responsible.

First, consider smoking as an alternative explanation. When one has a confounding factor that shows such great variation in prevalence between subsets of the population and which is associated at its highest levels of exposure with increases in the risk of the disease of interest with relative risks of 20–30 and higher, it is disingenuous to believe that the effect of this confounder has been eliminated with the measures taken in this study. There is indeed evidence in the study itself that this cohort smoked more heavily than the U.S. general population, in that the increased risk was halved by only the very crude adjustment that the available data permitted. The statement in the authors' abstract that "neither smoking nor geographic location fully explain the increased lung cancer risk" is totally unjustifiable. A more modest (and accurate) statement of the situation would have been "Neither smoking nor geographic location *as measured in this study* fully explain the increased lung cancer risk." Why was this adjustment too crude to evaluate the full impact of the cohort's smoking patterns?

1. I have already discussed in the

context of the comments on Steenland and Ward the impossibility of obtaining from data for a single year the cumulative experience with cigarette smoking of a cohort that is being observed over several decades and that is changing its size and cigarette habits over the study period, much less of comparing this experience with that of a general population in which the practices vary geographically and are undergoing marked changes over time.

2. The Public Health Survey from which the smoking patterns in the cohort derived included only 16% of the cohort and none from three of the plants, including Lorain, which showed the highest SMR of any plant and from which 40% (23) of the excess lung cancer deaths derived.

3. Although the methods of collection of smoking data are more likely to be similar in the compared surveys here than in the study of Steenland and Ward, there is still no assurance that comparable methods were used to elicit the information.

4. The full value of smoking habits as a predictor of lung cancer risk is not captured in the four crude categories of smokers used in this study. Apart from the possibility of variation in levels of use within these categories, other variables are extremely important in measuring the impact of smoking. These include age at which smoking was begun (perhaps the single most important determinant of the lung cancer risk associated with smoking), duration of smoking, fraction of each cigarette typically smoked, and other such factors. In a group that already manifests a high smoking frequency as measured by current smoking (as does the Ward et al. cohort), these other variables assume great importance. Sterling and Weinkam, using data from the Household Interview Survey conducted by the National Center for Health Statistics in 1970, showed that "craftsmen, foremen, operative, operatives and kindred workers" (of which this cohort would mainly consist) not only had higher smoking rates than professionals and managers, but began to smoke at earlier ages.[22] If the very crude adjustment for smoking used in this study accounts for half the excess lung cancer risk, I would have no difficulty believing that a proper adjustment would explain the entire excess.

While I do not rule out the possibility that the excess lung cancer observed in this cohort is attributable to occupational exposure to beryllium, there are a number of peculiar features of the data that do not seem to support that view. One of these is the lack of evidence of relationship between level or duration of exposure and lung cancer risk. There are no direct measures of exposure in either this or the Steenland and Ward study, but there is no evidence of any relationship when the usual and reasonable surrogates of level of exposure are used. Thus:

a. There is no relationship to duration of employment. The relationship to duration of employment (there are no data on duration of exposure) appears actually to be inverse (Steenland and Ward's Table VII), with only persons employed for less than 1 year showing a significant increase in lung cancer risk. This may be a function of heavy representation of Lorain employees among workers employed for less than 1 year; almost 20% of the employees with less than 1 year of employment worked at Lorain.

b. Around 1950 there was a considerable reduction in levels of exposure in these plants following the initiation of controls mandated by Atomic Energy Commission contracts in 1949–51. The trend in reduction in levels of exposure has continued since 1950. The concentration of beryllium in the air of extraction plants in the 1940s was commonly in excess of 1,000 $\mu g/m^3$, compared to the current 2 $\mu g/m^3$ OSHA standard that most plants now generally attain.[23] The authors have used these trends as a surrogate of exposure and show decreasing SMRs for workers first employed before 1950 (1.42), in the 1950s (1.24), and in the 1960s (0.62). However, these figures are seriously confounded by plant of employment. For example, Lorain, which shows the highest levels of lung cancer risk, was closed in the late 1940s, so had no experience of workers first hired after 1950; Lucky, Elmore, Hazelton, and St. Clair had no workers employed before 1950. The authors' Table XI shows that there were two plants with experience both of workers first employed before and of employees first hired after 1950 (Reading and Cleveland). In both of these plants, the SMRs were actually higher for workers hired in 1950–59 than for those hired before 1950. SMRs are unequivocally lower among workers first employed in 1960 or later (indeed, in the cohort as a whole, SMRs for such employees are well below 1.0, and significantly so). The overall pattern is not one that can readily be reconciled with the expected decrease in risk with increased control of exposure. The 12 cases in "multiple plants" (4.74 expected) hired before 1950 would seem to warrant some further investigation.

c. Ward et al. note that lung cancer was particularly elevated (SMR = 3.33, CI = 1.66–5.95) among workers at the Lorain plant with a history of acute disease—an observation that agrees with that of Steenland and Ward, which is not entirely surprising, since many of the cases must overlap the two studies. I have commented on the possible peculiarities of the Lorain plant in the context of Steenland and Ward and will return to it later.

In sum, the only clear-cut evidence of a relationship of lung cancer risk to level of exposure is the fact that lung cancer deaths in this study were *fewer* (18) than expected (28.9) (by a margin that is statistically significant) among workers first hired after 1959. Among the exposures characteristic of the period before 1959 (which comprises the bulk of the experience observed) there is no clear-cut relationship between level of exposure and level of risk if confounding by plant of employment is taken into account. This is a most unusual observation in the context of occupational carcinogenesis and is difficult to reconcile with the idea that the occupational exposure is responsible.

Conclusion Regarding the Post-1987 Studies

Although the authors of both of the post-1987 studies conclude that occupational exposure to beryllium is

the most likely explanation of the significant excesses of lung cancer deaths that they observe, their cases are not well founded. I note:

1. The very small size of the excesses observed, particularly in the study of Ward et al., where the SMR, after a very crude adjustment for smoking, is only 1.13. I am unaware of any known carcinogen that is associated with such a small elevation of risk in populations in which the exposure was fairly extensive. Mantel, paraphrasing one of the most important and frequently cited papers in the literature of statistical epidemiology,[24] states: "In Mantel and Haenszel, we had suggested that a minimum RR of 1.5 should be required so as to preclude any possible effect of an unrecognized bias."[25] While the original statement was made in the context of case-control studies, it seems no less applicable to cohort studies, particularly when, as in the present instance, potential confounders are measured with considerably less accuracy than in the typical case-control study.

2. Neither study takes adequate account of the possibility that the study cohorts had a higher prevalence of cigarette smoking than the general population. In this disease (lung cancer) cigarette smoking is an overwhelming potential confounder, small differences in which between compared populations can lead to important differences in risk. In the paper by Ward et al there is actual evidence that the cohort did have a higher prevalence of smoking than the general population, and other patterns of excess disease associated with smoking in this cohort (excess ischemic heart disease and emphysema) add to this evidence. Indications of excess smoking are not to be found in the paper by Steenland and Ward, but in this paper the comparison is of persons already suffering from respiratory disease (sufficiently severe to get them into the BCR) with the generally healthy population—a comparison that makes no sense whatsoever.

3. In neither paper is there firm evidence of increased risk of lung cancer associated with increased level of exposure, as measured by a number of surrogates. This would be most unusual for an occupational carcinogen.

4. The two papers produce inconsistent results with respect to whether or not the occurrence of lung cancer is associated with a period of latency after first exposure. It is admitted, however, that the information on duration of exposure in the paper by Steenland and Ward "is likely to contain numerous inaccuracies." (p. 1382)

5. The higher reported risk of lung cancer associated with the acute form of beryllium disease in both studies appears to be limited to the experience at Lorain, which closed in 1948. The centrality of this plant's experience to the findings in the cohort as a whole must be of great concern to those who wish to attribute the small excess risk in the whole cohort to occupational exposure to beryllium. It is true that exposures at Lorain were very high, but little is known about other relevant features of the processes in the plant or of the characteristics of its employees.

Conclusions from the Epidemiologic Evidence Overall

It appears to me that the studies published after 1987 add very little to the overall case for the carcinogenicity of beryllium in humans. While the studies are technically more sound than those of earlier years, the inferences that can be drawn from them are still weak, particularly because of the investigators' inability to deal adequately with the very likely presence of confounding by cigarette smoking over and above that which was measured. It is possible that beryllium, in much larger doses than occupational groups have been exposed to in the last 40 years, is in fact associated with carcinogenesis in the lung, but alternative explanations cannot be ruled out on the basis of available evidence. Even if such doses were carcinogenic, their relevance to occupational experience of the last three decades is nil. The situation exemplifies, perhaps, the futility, in the context of public health and public policy, of defining whether a chemical is or is not carcinogenic to humans without due consideration of dosages and routes and mechanisms by which humans are exposed.

Acknowledgment

This review was requested and funded by the Beryllium Industry Scientific Advisory Committee, an independent committee supported by the Brush Wellman and NGK Metals Companies. The opinions expressed are those of the author alone.

References

1. Tepper LB, Hardy HL, Chamberlin RI. *Toxicity of Beryllium Compounds*. Amsterdam: Elsevier; 1961.
2. Rossman MD, Preuss OP, Powers MB (Eds). *Beryllium. Biomedical and Environmental Aspects*. Baltimore, MD: Williams & Wilkins; 1991.
3. United States Environmental Protection Agency. *Health Assessment Document for Beryllium*. (EPA/600/8-84 026F). Research Triangle Park, NC: EPA; 1987.
4. International Agency for Research on Cancer. Some metals and metallic compounds. *IARC Monogr Carcinog Risk Chem Hum*. 1980;23.
5. International Agency for Research on Cancer. Overall evaluations of carcinogenicity: an updating of IARC monographs Volumes 1 to 42. *IARC Monogr Carcinog Risk Chem Hum*. 1987; Suppl. 7.
6. Steenland K, Ward E. Lung cancer incidence among patients with beryllium disease: a cohort mortality study. *J Natl Cancer Inst*. 1991;83:1380–1385.
7. Ward E, Okun A, Ruder A, Fingerhut M, Steenland K. A mortality study of workers at seven beryllium processing plants. *Am J Indust Med*. 1992;22: 885–904.
8. Mancuso TF, El-Attar AA. Epidemiologic study of the beryllium industry. Cohort methodology and mortality studies. *J Occup Med*. 1969;11:422–434.
9. Mancuso TF. Relation of duration of employment and prior respiratory disease among beryllium workers. *Environ Res*. 1970;3:251–275.
10. Mancuso TF. Occupational lung cancer among beryllium workers. In: Lemen R, Dement J, eds. *Dust and Diseases*. Forest Park, IL: Pathotox Pub; 1979:463–472.
11. Mancuso TF. Mortality study of beryllium industry workers' occupational lung cancer. *Environ Res*. 1980;21: 48–55.
12. Wagoner JK, Infante PF, Bayliss DL. Beryllium: an etiologic agent in the induction of lung cancer, nonneoplastic respiratory disease and heart disease among industrially exposed workers. *Environ Res*. 1980;21:15–34.
13. Bayliss DL, Lainhart WS, Crally LJ, Ligo R, Ayer H, Hunter F. Mortality patterns

in a group of former beryllium workers. In: *Trans 33rd Ann Meeting Am Conf Gov Indust Hygienists: Toronto, Canada*. Unpublished; 1971. pp 94–107.
14. Bayliss DL, Lainhart WH. Mortality patterns in beryllium production workers. Presented at American Industrial Hygiene Association Conference, May population 18, 1972. Unpublished. OSHA Administrative Law Beryllium Hearing; exhibit No. 66, docket No. H-005, 1977.
15. Bayliss DL, Wagoner JK. Bronchogenic cancer and cardiorespiratory disease mortality among white males employed in a beryllium production facility. Cincinnati, OH; NIOSH, Industry-wide Studies Branch. OSHA Beryllium Administrative Law Hearing; 1977. Exhibit 13.F.
16. Bayliss DL. Letter to William H. Foege, M.D., Director, Center for Disease Control. Atlanta, GA; unpublished, 1980.
17. Hammond EC. Smoking in relation to the death rates of one million men and women. *Natl Cancer Inst Monogr*. 1966; 19:127–204.
18. Millar JD. Memorandum to Director, CDC. Public Health Service, September 26, 1978.
19. Infante PF, Wagoner JK, Sprince NL. Mortality patterns from lung cancer and nonneoplastic respiratory disease among white males in the Beryllium Case Registry. *Environ Res*. 1980;21:35–43.
20. Eisenbud M. Lung cancer incidence among patients with beryllium disease (Letter to the editor). *J Natl Cancer Inst*. 1993;85:1697–1698.
21. Steenland K, Ward E. Lung cancer incidence among patients with beryllium disease (Letter to the editor). *J Natl Cancer Inst*. 1993;85:1698–1699.
22. Sterling TD, Weinkam JJ. Smoking characteristics by type of employment. *J Occup Med*. 1976;18:743–752.
23. Eisenbud M, Lisson J. Epidemiologic aspects of beryllium-induced nonmalignant lung disease: a 30-year update. *J Occup Med*. 1983;25:196–202.
24. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst*. 1959;22:719–748.
25. Mantel N. Active and passive smoking and pathological indicators of lung cancer—a report of limited value? *JAMA*. 1993;270:1689–1690.