



Strom DJ. A Critique of "Mortality Among Workers at Oak Ridge National Laboratory." *Nuclear News* 34(9):67-74, July 1991.

# A critique of "Mortality Among Workers at Oak Ridge National Laboratory"

by Daniel J. Strom

A low-dose radiation epidemiology paper titled "Mortality Among Workers at Oak Ridge National Laboratory" appeared in the March 20, 1991, issue of the *Journal of the American Medical Association*.<sup>1</sup> The authors, S. Wing et al., are funded by the U.S. Department of Energy (DOE) and work either at the University of North Carolina (UNC), at Oak Ridge Associated Universities (ORAU), or at Oak Ridge National Laboratory (ORNL).

The paper made some startling claims and attracted media attention. It contains an attack on uses of radiation, and makes sweeping statements regarding public health and radiation risks. In this article, I examine the Wing et al. paper in light of the problems of low-dose epidemiology as noted by the ICRP, UNSCEAR, and the BEIR V Committee and in Sir Austin Bradford Hill's criteria for inference of causality, and add some thoughts of my own. (Throughout this article, as in the paper being reviewed, "dose" is used for "dose equivalent.")

## Low-dose radiation epidemiology

A "low dose of radiation" is generally taken to be below 250 milligrays (mGy; 25 rads) or 250 millisieverts (mSv; 25 rem) for low-linear energy transfer (LET) radiations such as gammas and betas.<sup>2</sup> Radiation risk estimates by groups such as the International Commission on Radiological Protection (ICRP), the United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR), and the National Academy of Sciences' National Research Council Committee on the Biological Effects of Ionizing Radiation (BEIR Committee) are generally based on extrapolations from "high" doses.

Daniel J. Strom, Ph.D. and certified health physicist, is an associate professor of Environmental and Occupational Health, Graduate School of Public Health, at the University of Pittsburgh.

TABLE I  
FACTORS TO CONSIDER BEFORE INFERRING THAT AN ASSOCIATION IS A CAUSAL ONE (FROM REF. 6)

1. Strength of the association: Is a large effect seen—for example, a 32-fold lung cancer increase in heavy smokers?
2. Consistency: Is the effect consistently observed across studies?
3. Specificity: Are particular sites and types of effect associated with exposure?
4. Temporality: Does exposure precede the effect?
5. Biological gradient: Is there a dose-response relationship?
6. Plausibility: Is it biologically plausible that this exposure could cause this effect? (Biological plausibility depends to some extent on how much biology one knows.)
7. Coherence: Does this association seriously conflict with generally known facts of the natural history and biology of the effect?
8. Experiment: Does intervention to prevent exposure reduce or prevent the effect?
9. Analogy: Do other, similar agents produce the effect?

These scientific groups all discuss the limitations of the low-dose studies, and explain why they use high-dose studies with extrapolations.

The ICRP in its 1990 Recommendations discussed "reports involving exposure of populations to low doses of radiation."<sup>3</sup> The ICRP stated that "... these studies suffer from one or more of the following methodological difficulties including small sample size, lack of adequate controls, extraneous effects other than those due to radiation, inadequate dosimetry and confounding social factors. Furthermore 'positive' findings tend to be reported while negative studies often are not. Overall, studies at low dose, while potentially highly relevant to the radiation protection problem, have contributed little to quantitative estimates of risk."

The 1988 UNSCEAR Report<sup>4</sup> also discussed the limitations of low-dose studies. It stated that "an assessment of effects of low dose is clouded by the need for large samples, the difficulty of accurately estimating exposure and the growing importance of extraneous sources of variation, including diagnostic and therapeutic exposures that are less compromising when the doses are large."

The BEIR Committee in its 1990 BEIR V Report devoted several pages to low-dose studies.<sup>5</sup> The BEIR Committee

pointed out the problems of random error and various kinds of bias, including selection bias, information bias, confounding, the ecological fallacy, publication bias, and multiple comparisons.

Carefully designed and executed epidemiology studies can show statistical associations between disease and exposure. But showing association does not necessarily show causation—that is, that the exposure caused the disease. My favorite example of a noncausal association is the finding that "In the winter I wear galoshes. In the winter, I have colds. Therefore, galoshes cause colds." A statistician could show that the association between galoshes and colds is fairly strong, but for one to infer that galoshes cause colds from this association requires a great leap of faith.

There are several factors to consider before inferring that an association is a causal one, as shown in Table I. These include the strength of the association, consistency, specificity, temporality, biological gradient, plausibility, coherence, intervention, and analogy.<sup>6</sup> Clearly the galoshes-colds association does not meet many of these factors. We will look at these for the ORNL study.

## The UNC/ORAU/ORNL findings

The study listed 17 517 workers employed at ORNL between 1943 and 1972,

and focused on a subcohort of 8318 white males. Among that group, 1524 had died by 1984, the most recent year of follow-up.

Some of the study results are expressed as "standardized mortality ratios" (SMRs), which compare death rates within the cohort to death rates that would be expected in a similar group of U.S. white males of similar ages. Since death rates increase dramatically with age, the adjustment for ages is required for any meaningful comparisons. This adjustment is called "standardizing." SMRs can be calculated for all causes of death, from all cancers, from heart disease, from leukemia, etc.

In their reports on the study, the news media did not focus on the fact that the all-cause SMR was 0.74, that is, ORNL

males, a 63 percent excess (which should happen by chance less than one time in 20).

Furthermore, among a subgroup of "workers monitored for internal contamination," a 123 percent leukemia excess was seen (16 cases where just over 7 were expected), along with a 65 percent excess of lymphosarcoma and reticulosarcoma (6 cases where 3.6 were expected). Nine other cancer types, however, including 127 of the 149 cancer deaths in the "workers monitored for internal contamination" group, occurred much less frequently than expected, resulting in an all-cancer SMR of 0.82.

Despite the low death rate from all cancers combined, the excesses in leukemia, a rare cancer, may be cause for concern, since leukemia (other than

In view of the limitations of low-dose epidemiology, the Wing et al. paper has several strengths and many weaknesses.

### Strengths of the study

This is generally a well-designed study with individual (rather than group) dose measurements and individual causes of death. The study's strengths include the fact that it covers a fairly large, well-defined cohort with many years of follow-up and 91.8 percent vital status known. Thus, only 8.2 percent of the workers were untraceable, a low rate that is comforting. Furthermore, there were actual measured external doses for many of these workers from 1942 through 1984, using pocket ion chambers, film, and thermoluminescent dosimeters. Note that each exposure measurement is associated with an individual, a phenomenon almost unheard of outside of occupational radiation epidemiology. Even the Japanese bomb survivors have doses calculated, not measured, for each person based on where he or she was at the time of the blast. Having individual measurements makes the dosimetry information particularly robust. The dose variables for each individual were carefully compiled by health physicists, some of whose names appear in the acknowledgments.

The use of internal comparisons is a good design feature. The analyses of dose-response relationships does not involve an external comparison group such as U.S. white males, and shows only how workers at ORNL compare to each other. This kind of internal comparison usually eliminates many problems.

### Weaknesses of the study

There are enough weaknesses in this study and studies like it, however, that it is hard to justify all the attention they receive. The list of problems is given in Table II.

The ORNL study meets many of Hill's criteria for inference of causality, but all findings fail to meet one or more criteria. The study suffers from most of the shortcomings that the ICRP, NAS, and United Nations committees discussed. Furthermore, the interpretation of the findings and the antinuclear diatribe range from not very scientific to simply wrong.

Wing et al.'s insistence on expressing their opinions on radiation and nuclear power resulted in some prominent co-workers' removing their names from authorship (you'll find them in the acknowledgments). Consider, for instance, the following paragraph from the paper:

Conversely, while factors other than radiation clearly predominate the statistical analysis of mortality in this popula-

---

***In their reports on the study, the news media did not focus on the fact that . . . ORNL workers die only 74 percent as fast as U.S. white males. . . . For every three 70-year-olds from the Oak Ridge cohort who died, about four would have died in a group of U.S. white males of the same size.***

---

workers die only 74 percent as fast as U.S. white males. To make this more concrete, consider a group of 70-year-olds. For every three 70-year-olds from the Oak Ridge cohort who died, about four would have died in a group of U.S. white males of the same size. Thus, ORNL workers were and are a very healthy bunch of people. The all-cancer SMR was 0.79, meaning ORNL workers had only 79 percent of the cancer death rate of U.S. white males.

Most working populations are healthier than the average U.S. white male, since the group of all U.S. white males includes people who are too sick to work. This widely known phenomenon is called the "healthy worker effect" and is not unexpected. These two SMRs, however, are quite low even for working populations.

So what's all the fuss about? Researchers generally want to look at their data in as many ways as possible, so they look at finer and finer subdivisions of disease categories. For the disease category "all leukemias" (including chronic lymphocytic leukemia), the ORNL study found 28 cases when they would have expected about 17 in a similar group of U.S. white

chronic lymphocytic leukemia) has been seen in excess in populations exposed to high doses of radiation. For this reason, the authors chose to find out whether there was a relationship between radiation doses and cancer risk within the ORNL cohort. When an excess (or deficit) of some cause of death is observed, a researcher has more confidence that the dose caused the excess risk if there is more excess (or deficit) when there is more dose. The inference that a statistical association is causal, that is, that radiation causes excess cancer, is strengthened if a "dose-response relationship" can be found (Hill's criterion 5 in Table I). On the other hand, if no dose-response relationship is seen, then confidence in the findings is weakened.

Wing and his co-workers did find statistically significant, positive dose-response relationships in the ORNL workers for all cancers and for a couple of other selected endpoints, but not for leukemia. The strength of the relationship for all cancers (when doses within 20 years of death were ignored) is what attracted the attention, and led to claims that radiation is 10 times more hazardous than previously reported.

tion, the public health impact of these radiation exposures and the industry that produces them [i.e., the nuclear industry] extend far beyond the low-dose occupational exposures themselves, which are estimated to constitute only 0.3% of the population dose of ionizing radiation in the United States. The exposure of workers in this setting, and any attending health effects, depends on the historical development of an industry linked to a concentration of resources in military spending, which itself has gross health effects. By providing an alternative to fossil fuels for electric power generation, the [nuclear] industry encourages ever-increasing energy consumption, a factor in potential health effects of global climatic and environmental change. Additional effects of ionizing radiation from the industry may occur in surrounding communities, in offspring of workers, and in areas where waste products must be isolated from the environment for generations. Use of radioisotopes in medical research, diagnosis, and therapy also affects public health. Further consideration of potential harmful effects of low-dose ionizing radiation is essential in setting occupational and environmental exposure standards. However, focus on these effects should not distract attention from the public health impact of the context in which the exposures occur. [Brackets mine.]

It is amazing to me that these opinions within the context of a scientific paper survived the peer review process.

There are significant problems with random error due to sample size. Many cancer categories had fewer than 10 deaths in them, some with only 1 or 2. When the dose-response analyses were conducted, many dose groupings had no cases in them; these relationships are based on only a very few cases. Thus, the strength of the associations, while statistically significant for some of the few carefully chosen examples in the paper, is weakened by the small numbers.

The dose-response relationship for lung cancer is based on a few cases at high doses, with a very large deficit in total cases (SMR = 0.65!). One epidemiologist I spoke with stated that the risk at low doses was so low that the people were "practically immortal" as far as lung cancer was concerned.

Extraneous effects other than those due to radiation were certainly present at ORNL. Some of these may be associated with dose, making them confounders. For example, this study provides no data on chemical exposures.

The radiation dosimetry in this study is about as good as it gets in occupational studies, but in my view still suffers from major flaws. I regret that the collective dose equivalent (the sum of all doses) was not reported, since this would allow quick comparisons of risk with published

studies. The footnote to Wing et al.'s Table 2 regarding mean, median, and total doses makes no sense; there were no health physicists among the authors.

Human beings get doses from natural background, medicine, and, for studies such as this, from three occupational sources: work at ORNL; work at other DOE or predecessor agencies such as the Manhattan Engineer District, the Atomic Energy Commission, or the Energy Research and Development Administration; and work at non-DOE facilities. Since authors had access to

TABLE II  
PROBLEMS WITH THE  
ORNL MORTALITY STUDY

PROBLEMS IN DESIGN	
Small numbers of cases cause large statistical uncertainties	
Large unrecorded radiation exposures before the Manhattan Project, and unrecorded doses at non-DOE facilities	
Inclusion of high doses in a low-dose study	
Multiple comparisons not guided by prior hypotheses result in spurious correlations	
PROBLEMS IN INTERPRETATION OR MEDIA REPORTING	
Ignores the finding that this is a very healthy population (74 percent all-cause SMR)	
Effect of socioeconomic status (SES) is very large	
No dose-response relationship for leukemia	
Dose-response for lung cancer based on very few cases	
Selective reporting	
Not a consensus document	
Included "contributing" cancers	
Publication bias	
This is only one study	

work records at other DOE facilities, 3707 workers who had been at other DOE plants were eliminated from the study cohort.

A serious problem with the dosimetry remains, however, and that is the problem of large, unrecorded radiation doses that occurred in many of these people prior to 1942. As an example, consider Enrico Fermi, who is doubtless in this cohort. He and many of the others who worked in nuclear physics in the 1920s and 1930s received massive doses that were simply not monitored before 1942. There are many anecdotes in the literature concerning radiation-induced cat-

aracts and lowered white cell counts, both *high*-dose phenomena. In 1983, I recommended to the ORAU/UNC collaboration that these persons (I listed dozens) be flagged for special study, but there is no mention of this in the Wing et al. paper. Furthermore, no account is taken of natural background exposures, medical exposures, or exposures at non-DOE facilities, which may be confounders. Any excess cases of cancer in this study are being attributed to dose numbers that are surely underestimates, leading to inflated risk estimates.

Other dosimetry problems include the fact that neutrons were not well monitored in early days. Internal dosimetry is essentially unavailable. There is no account in the refereed literature of the conversion of dose records made for radiation protection purposes to doses suitable for epidemiology.

By far the strongest association found in this study was that with socioeconomic status (SES) as represented by pay code (hourly, weekly, monthly). It has long been known that educated people with good incomes generally take better care of themselves and get better medical care than poor, uneducated people. Countless studies show that such professionals as PhD researchers (who were paid monthly at ORNL) live longer than laborers (paid hourly) or even skilled craftsmen (paid weekly). In the presence of these huge effects, finding a radiation dose-response relationship is difficult. For example, for workers born after 1915 and paid monthly, the all-cancer mortality rate ratio was about 50 percent! (This information is found in the supplementary work not in the JAMA paper but referenced therein.)

Some workers received doses above the National Council on Radiation Protection and Measurements' 250-mSv cutoff, although at low dose rates. There were 19 workers with lifetime doses above 500 mSv (50 rem). The two individuals with doses greater than 1 Sv (100 rem) are still alive. Effects in these workers are hard to generalize to the lower doses seen today in the nuclear industry.

There are four particularly troubling aspects of the study design that affect its conclusions. First, the authors chose to include cancers that were "contributing causes" of death with cancers that were "underlying causes." This was a mortality study, where death certificates were examined to determine cause of death. Death certificates, however, have space for three conceptually different causes of death: immediate, underlying, and contributing causes. For example, many people die of pneumonia, commonly listed as the immediate cause of death. The underlying cause may be chronic obstructive pulmonary disease or lung

cancer. A contributing cause may be anemia. Deaths are usually ascribed to the underlying cause. The addition of contributing cancers makes it impossible to compare this study's findings directly with those of the BEIR Committee, which uses different criteria. The addition of contributing cancers essentially inflates the risk estimates.

Second, the researchers included chronic lymphocytic leukemia (CLL) in their leukemia category. This fails Hill's consistency and specificity tests, since CLL has never been associated with radiation in many larger, more powerful studies. We all want to know what the leukemia SMRs would be if CLL is excluded, but they don't tell us.

Third, this study as presented suffers from a problem of multiple comparisons not based on prior hypotheses. It has been said that data are like a captured spy: If you torture them long enough, they will tell you anything that you want to know. The BEIR V Committee is particularly emphatic about presenting analyses not guided by hypothesis (BEIR V, pp. 48-49). If you test 20 relationships at a 95 percent confidence level, one is likely to come out as significant simply by chance. If you test 200, then 10 will be significant.

Finally, there is the problem of selective reporting of results. Were there dose-dependent decreases in cancer deaths? We are told that most cancers were rarer among ORNL workers than among U.S. white males; however, only those for which there was excess are analyzed for dose-response effects. Many results are not reported in the JAMA paper, but the reader is referred to a supplementary, nonrefereed paper that contains less alarming results. Despite the authors' claims about hypothesis testing, the ORNL study's comparisons with different years of dose lagging for latency, with or without internal contamination monitoring, and so on, give the impression that only the tests that came out positive were reported. By their own admission, only the "best-fitting" lag assumptions were printed.

Both the ICRP and the BEIR committee noted the selective publication of positive results. No excess cancer is no news; professors do not get promotions, tenure, funding, or fame from well-designed studies that show no effect. The BEIR Committee stated that "epidemiologists are more likely to report and journal editors are more likely to accept positive findings than null findings." This is called publication bias, and has recently been recognized as a real problem in the medical literature.<sup>7</sup> The Wing et al. paper seems to be proof of publication bias as far as the news media are concerned; other studies with null

findings by many of the same researchers received little or no media attention.

There is one last caveat. The Wing et al. paper is, after all, only one study. It is not consistent with other studies, and at best adds only a small amount to our knowledge of radiation effects in people.

### Conclusions

The ORNL workers followed through 1984 were a very healthy group overall when compared to external groups. Most of the internal comparisons showed a striking effect of pay code, a surrogate for socioeconomic status. The results are compromised by design and methodological issues such as small numbers of cases, incomplete assessment of radiation doses from before 1942 and at other sites, inclusion of high-dose workers, selective reporting of results, inclusion of chronic lymphocytic leukemia, absence of a dose-response relationship for leukemia, and multiple comparisons not guided by prior hypotheses. Disagreements among the authors, selective reporting of results, inclusion of contributing cancers with underlying cancers, and the authors' own biases further weaken the study's credibility.

The admonitions of the International Commission on Radiological Protection, the National Academy of Sciences, and the United Nations Scientific Committee on the Effects of Atomic Radiation all lead us to treat the findings of Wing and co-workers with great caution. Plans to combine the occupational studies from Hanford, ORNL, and Rocky Flats plants, with British and Canadian facilities, along with more years of follow-up, should address some of these problems.

### References

1. S. Wing, C. M. Shy, J. L. Wood, S. Wolf, D. L. Cragle, and E. L. Frome, "Mortality Among Workers at Oak Ridge National Laboratory," *J. Am. Medical Assoc.*, **265**, 11, 1397-1402 (Mar. 20, 1991).
2. National Council on Radiation Protection and Measurements, *Influence of Dose and Its Distribution in Time on Dose-Response Relationships for Low-LET Radiations*, NCRP Report No. 64, Bethesda, Md. (1988).
3. International Commission on Radiological Protection, *1990 Recommendations of the International Commission on Radiological Protection*, ICRP Publication 60, Oxford: Pergamon Press (1991).
4. United Nations Scientific Committee on the Effects of Atomic Radiation, *Sources, Effects and Risks of Ionizing Radiation* (the 1988 UNSCEAR Report), Albany, N.Y.: Unipub (1988).
5. National Academy of Sciences, Committee on the Biological Effects of Ionizing Radiations, *Health Effects of Exposure to Low Levels of Ionizing Radiation: BEIR V*, Washington, D.C.: National Academy Press (1990).
6. A. B. Hill, "The Environment and Disease: Association of Causation?" *Proc. Roy. Soc. Med.*, **58**, 295-300 (1965).
7. C. B. Begg and J. A. Berlin, "Publication Bias: A Problem in Interpreting Medical Data," *J. Roy. Statist. Soc. A*, **151**, 3, 419-463 (1988).