metric introduced there or the one defined by (1) is considered. The distance (1) is called the *distance with respect to the mapping* (s, F) by L. Cesari, who introduces it in compact metric spaces (cf. "Surface Area," *Ann. Math. Studies*, **35**, 159 (1956).

[5] Cf. RS, §7(f).

[6] Cf. *ibid.*, §3(c).

[7] Cf. C. Kuratowski, *Topologie* (Warzawa-Wrocław, 1950), II, 51, théor. 4.

[8] Cf. *ibid.*, I, 176, théor. 1; cf. also Hurewicz-Wallman, *Dimension Theory* (Princeton: Princeton University-Press, 1948), p. 30, theor. III 2.

[9] Cf. W. Gross, "Eine ganze Funktion, für die jede komplexe Zahl Konvergenzwert ist," *Math. Ann.*, **79**, 201–208, 1918.

[10] Cf. RS, §4 for the relationship between asymptotic paths and boundary points and §5 for the notation used here.

[11] Cf. ibid. §3(d).

[12] $L_{-1} = \phi$.

[13] Apparently every boundary point is the limit of a sequence of branch points. Indeed, if $\zeta \epsilon z$, either every neighborhood of ζ contains parts of infinitely many sheets, or all sufficiently small neighborhoods are contained in a single sheet F_n. In the first case ζ is a cluster point of cuts $L_{n,\mu}$ and therefore, according to (ii), also of branch points. The second case cannot occur because every boundary point of a sheet F_n is an interior point of F.

[14] We can, of course, assume A to be non-empty.

---

*THE EFFECTS OF STRONTIUM-90 ON MICE**

BY BARCLAY KAMB AND LINUS PAULING

CALIFORNIA INSTITUTE OF TECHNOLOGY, PASADENA, CALIFORNIA

Communicated November 17, 1958

On Sept. 19, 1958 there was published in *Science* a paper by Dr. Miriam P. Finkel of Argonne National Laboratory in which she communicated her observations on the effects of strontium-90 injected into mice on life expectancy and on incidence of tumors of bone and blood-forming tissues.[1] She discussed the question of whether or not the effects are proportional to the amount of injected strontium-90 at low doses, and reached the conclusion that it is likely that there is a threshold with value for man between 5 and 15 μc. (as compared with the present average value from fallout, about 0.0002 μc., and the predicted steady-state value from fallout for testing of nuclear weapons at the average rate for the past five years, about 0.02 μc.). Her paper ends with the sentence "In any case, the present contamination with strontium-90 from fallout is so very much lower than any of these levels that it is extremely unlikely to induce even one bone tumor or one case of leukemia."

On the same day, Sept. 19, 1958, newspapers throughout the United States published accounts of this work. For example, the Pasadena (Calif.) *Star-News* contained an article with the headline "Tests on Mice Show Fallout Safe" and the first sentence, "A woman researcher says tests on mice show that the present fallout from nuclear weapons tests will not produce a single case of bone cancer or leukemia in humans." The *New York Times* published accounts of the work on both September 19 and September 28.

We have made an analysis of Dr. Finkel's data that shows that she had no justi-

fication whatever for her concluding statement. All of her data are compatible with a zero threshold for strontium-90. Moreover, the statistical analysis shows that in order for Dr. Finkel to have been justified with 90 per cent confidence (10 per cent type-II error) in making her concluding statement on the basis of her data she would have to have used over 1,000,000,000,000 mice in each of her groups, instead of the 150 or less that were used. It is hard for us to understand how such a serious error could be made in Dr. Finkel's argument, leading her to publish her seriously misleading statement about this matter of great importance.

The Mice Experiments.—In the studies described by Dr. Finkel young adult female mice (strain CF No. 1, about 70 days old) were given a single intravenous injection of an isotonic equilibrium mixture of strontium-90 chloride and yttrium-90 chloride. There were twelve injected groups, ranging in size from 150 mice for the group receiving the smallest amount (1.3 μc/kg body weight) to 15 for that receiving the largest amount (9330 μc/kg, an amount that caused death of about 50 per cent within 30 days). The control group contained 150 mice. The author states that there is 11 per cent retention (at 600 days) of the injected radioactive material. Report was made of the fractional decrease in average survival time, the incidence of animals with osteogenic sarcomas (among 150-day survivors), and the fractional decrease in time to a 20 per cent incidence of reticular tissue tumors compared with the 20 per cent incidence time of the controls.

Studies of this sort may be of great value in providing information about the probable amount of damage done to human beings by exposure to high-energy radiation, such as that from strontium-90 produced by nuclear weapons. It is important that the analysis of the experimental results be carried out correctly. We have found that in the treatment of problems of this sort the assumption that the probability of damage is strictly proportional to the amount of radiation exposure does not in general require that a response such as decrease in life expectancy be linear, except over a very small range. Moreover, we have found that this assumption together with the theory of statistics can be applied in a reasonably straightforward way in the discussion of data such as those obtained by Dr. Finkel, as shown in the following sections.

Analysis of the Experimental Data on Life Shortening.—Our analysis proceeds from the hypothesis, induced by Lewis[2] as a result of his study of the incidence of leukemia, that exposure of the bone marrow of an animal to radiation results in an increase in the probability per unit time that the animal will die at any time thereafter, the increase being proportional to the quantity of radiation absorbed. We shall suppose that this hypothesis applies to all of the radiation-induced effects in Dr. Finkel's experiments with mice.

Let N_0 be the number of animals at the beginning of a given experiment, taken to be at $t = 0$, and let $N(t)$ be the expected number (average for many experiments of the same kind) at the later time t. Further, let $N^0(t)$ be the expected number in a "control" experiment in which no strontium-90 is injected, so that

$$g(t) \equiv -\frac{1}{N_0}\frac{dN^0}{dt}$$

is the natural specific death-rate function. We denote by α the quantity of strontium-90, in μc/kg body weight, that is retained in the animals. Then our hypothesis yields the equation

$$\frac{dN}{dt} = -N\beta\alpha t - \frac{N}{N^0(t)} \cdot N_0 g(t) \tag{1}$$

where β is a constant of proportionality relating the quantity of strontium-90 retained to the increased probability per unit time that the animals will die, this probability of course increasing linearly with time owing to the nearly constant irradiation by the decaying strontium-90.

Let $n^0(t) = N^0(t)/N_0$, so that $g(t) = -\dot{n}^0(t)$. Then on integrating equation (1) we obtain

$$\begin{aligned} N &= N_0 \exp\left[-\frac{1}{2}\beta\alpha t^2 - \int_0^t \frac{\dot{n}^0(t)}{n^0(t)}\,dt\right] \\ &= N_0 n^0(t) e^{-(1/2)\beta\alpha t^2} = N^0(t) e^{-(1/2)\beta\alpha t^2} \end{aligned} \tag{2}$$

To compare this result with the experimental data we calculate Δ, the fractional decrease in life expectancy (fractional decrease in average survival time after injection),

$$\Delta = \frac{t_0 - t_\alpha}{t_0}$$

where t_α is the life expectancy for a retained quantity α of strontium-90,

$$t_\alpha = \frac{1}{N_0}\int_0^\infty N(t)dt \tag{3}$$

The equation for Δ is

$$\Delta = 1 - \frac{1}{t_0}\int_0^\infty n^0(t) e^{-\gamma t^2}\,dt \tag{4a}$$

$$= \frac{1}{t_0}\int_0^\infty \left(t - \frac{1}{2}\sqrt{\frac{\pi}{\gamma}}\,\mathrm{Erf}\, t\sqrt{\gamma}\right) g(t)\,dt \tag{4b}$$

where for simplicity we put $\gamma = \frac{1}{2}\alpha\beta$. The result in equation (4b) is obtained by an integration by parts, and the error function is defined as

$$\mathrm{Erf}\, x = \frac{2}{\sqrt{\pi}}\int_0^x e^{-y^2}\,dy$$

If normally (for $\alpha = 0$) all animals lived to the age t_0 and then died, so that $g(t)$ were a delta function $\delta(t - t_0)$, then we would have simply

$$\Delta = 1 - \frac{\sqrt{\pi}}{2}\frac{\mathrm{Erf}\, t_0\sqrt{\gamma}}{t_0\sqrt{\gamma}} \tag{5}$$

However, the actual lifetimes scatter with sizable dispersion about t_0. The extent of this dispersion can be estimated from the acceptance region $\Delta \leq 0.07$ quoted by Dr. Finkel (Figs. 3 and 4) as appropriate to a test of the hypothesis of no difference between the responses of the control population and of a population injected with a given dose of strontium-90. If we assume (1) that the test of no difference applies to Curve A, the life-shortening data, (2) that the test was one-tailed, and (3) that the test accounted for the uncertainty in the mean lifetime of the control population and for the uncertainty in the mean lifetime of her group 10, the highest-dosed population to fall within the acceptance region (except for the "peculiar result" for group 8), then we find that the estimated standard deviation for $g(t)$ is $\hat{\sigma} = 258$ days.

These assumptions are somewhat uncertain, as explained later, but they are the best that can be made from the information given in Dr. Finkel's paper. The uncertainty in drawing any conclusions about $g(t)$ from Dr. Finkel's data lead us to take a more general approach. Gompertz discovered that for animal populations the logarithm of the "age-specific death rate" is closely a linearly increasing function of time. For man the age-specific death-rate doubling time is about 8 years. Jones[3] has pointed out that the doubling times for different animal species are approximately proportional to the mean life spans for the species. We shall use this information to derive a hypothetical death-rate function $g(t)$ for the mouse population used in Dr. Finkel's experiments.

The Gompertz law is

$$\ln g(t) = C + Bt \tag{6}$$

which yields

$$n^0(t) = \exp\,[-A(e^{Bt} - 1)] \tag{7}$$

where A $(= e^C/B)$ is a constant and where B is related to the doubling time τ_D by

$$B = \frac{\ln 2}{\tau_D} \tag{8}$$

A is to be chosen so as to give the correct mean life span:

$$t_0 = \int_0^\infty \exp\,[-A(e^{Bt} - 1)]dt = \frac{e^A}{B}\,[-\mathrm{Ei}(-A)] \tag{9}$$

The exponential integral $\mathrm{Ei}(x)$ is defined by Jahnke and Emde.[4] If $t = 0$ is taken to be a time shortly after birth, but long enough after birth to exclude infant deaths (which are omitted in Gompertz' treatment), then t_0 is T, the mean life span from birth to death. If then T/τ_D is a constant for all animal species, we find from equations (8) and (9) that A is a constant, independent of species, given by the solution of the equation

$$e^A[-\mathrm{Ei}(-A)] = \frac{T}{\tau_D} \ln 2 \tag{10}$$

Assuming $T = 60$ years for man, with $\tau_D = 8$ years, we find $A = 0.0032$. The solution of equation (9) is obtained with the help of the expansion[4]

$$-\mathrm{Ei}(-x) = -\ln\Gamma x + x - \frac{x^2}{2!2} + \ldots \tag{11}$$

where $\Gamma = 1.781$.

The death-rate function is

$$g(t) = AB \exp\left[Bt - A(e^{Bt} - 1)\right]$$

The dispersion of life spans is measured by the standard deviation σ of $g(t)$:

$$\sigma^2 = \int_0^\infty t^2 g(t)dt - t_0^2 = 2\int_0^\infty t n^0(t)dt - t_0^2$$

The second expression results from an integration by parts. A numerical integration is required to obtain σ^2, which can most easily be carried out with $n^0(t)$ values from equation (7). In this way we find from equations (9) and (10), with $t_0 = T$, that

$$\sigma = 0.24T$$

For Dr. Finkel's mice, reported to have $T = 670$ days, we have $\sigma = 161$ days, in rather poor agreement with the value $\sigma = 258$ days inferred above from her paper.

In the calculations that follow we have used equations (7) and (10), with the assumption $T = t_0 = 600$ days, although actually the mice were about 70 days old at the beginning of the experiment. Thus we have used a doubling time τ_D of 80 days, and our $g(t)$ has standard deviation $\sigma = 141$ days. The assumption $\tau_D = 80$ days agrees with the value quoted by Jones[3] for mice. Values of $n^0(t)$ for these parameters are given in Table 1. The difference between assuming $T = 600$ days and assuming $T = 670$ days is not great; in fact, survival curves calculated from equation (5), which assumes $\sigma = 0$, do not differ greatly from curves obtained by the more refined procedure that we have used.

TABLE 1

t (Days)	$n^0(t)$	$e^{-\mu t^2}$
0	1.000	1.000
80	0.997	0.997
160	0.990	0.987
240	0.978	0.972
320	0.953	0.950
400	0.908	0.923
480	0.818	0.891
560	0.666	...
640	0.443	...
720	0.195	...
800	0.038	...
880	0.000	...

We proceed now to compare equation (4a), evaluated with the help of equations (7) and (10), with the experimental life-shortening data. We assume with Dr. Finkel that $\alpha = 0.11\alpha^*$ where α^* is the injected dose of strontium-90 in μc/kg body weight, and we attempt to choose the available parameters so as best to reproduce the observed life-shortening data $\Delta(\alpha^*)$. There are two parameters: the constant β, and the no-dose life-shortening Δ_0, the latter arising from the fact that we cannot give great weight to Dr. Finkel's zero point because of the statistical uncertainty

in the observed mean life span of the control population. Thus the theoretical curve to be fitted to the data is

$$\Delta(\alpha^*) = \Delta_0 + 1 - \frac{1}{t_0}\int_0^\infty n^0(t)e^{-(0.11/2)\beta\alpha^* t^2}\,dt \tag{12}$$

The inclusion of Δ_0 simply as a constant in equation (12) is not strictly correct: the additive term arising from an adjustment of the zero point should be written $\Delta_0(\alpha^*)$, where $\Delta_0(\alpha^*)$ is a somewhat complicated decreasing function of α^* that tends to zero as $\alpha^* \to \infty$. For simplicity, however, we ignore this complication, which proves to be unimportant for the lower radiation levels ($\alpha^* \lesssim 1000$ μc/kg), and which in any case does not much change the results obtained, because Δ_0 is small.

For very small values of α^* equation (12) reduces to

$$\begin{aligned}\Delta(\alpha^*) &= \Delta_0 + \alpha^*\left[\frac{1}{2}(0.11)\beta\int_0^\infty n^0(t)t^2dt\right] \\ &= \Delta_0 + \alpha^*\cdot\frac{d\Delta}{d\alpha^*}\bigg|_{\alpha^*=0}\end{aligned} \tag{13}$$

This is the linear response region. We can therefore choose preliminary values of Δ_0 and β by estimating a linear fit to the experimental points at low values of α^*. The integration in equation (13) is performed numerically, with use of equation (7).

We have calculated theoretical curves from equation (12) in three steps: (1) For $\alpha^* \leq 50$ μc/kg equation (13) applies; (2) for selected values of α^* in the range $50 \leq \alpha^* \leq 1000$ we carry out the integration in equation (12) numerically, using time intervals $\Delta t = 80$ days; (3) for $\alpha^* > 1000$ it is found that the asymptotic form of equation (12) is valid:

$$\Delta(\alpha^*) = \Delta_0 + 1 - \frac{1}{t_0}\sqrt{\frac{\pi}{0.11\beta\alpha^*}} \tag{14}$$

Because we wish to examine the result statistically, we adjust the parameters by a weighted least squares procedure. We calculate two theoretical curves $y = f_0 + f(x, \beta_1)$ and $y = f_0 + f(x, \beta_2)$ (here y is fractional life shortening, Δ; x is injected dose, α^*; and f_0 is the constant Δ_0) for two nearly correct values β_1 and β_2 of the parameter β. We ask for values $\hat{f}_0 = f_0 + \Delta f_0$ and $\hat{\beta} = \beta_1 + (\beta_2 - \beta_1)\delta$ of the parameters such that the weighted sum of the squares of the differences between the experimental values y_i and the theoretical values $y(x_i)$ is a minimum:

$$\sum_i w_i(y_i - f(x_i, \hat{\beta}) - \hat{f}_0)^2 = \min \tag{15}$$

Since $\beta_2 - \beta_1$ is small we can assume that

$$\begin{aligned}f(x_i, \hat{\beta}) &= f(x_i, \beta_1) + [f(x_i, \beta_2) - f(x_i, \beta_1)]\delta \\ &\equiv f(x_i, \beta_1) + \Delta f_i\delta\end{aligned} \tag{16}$$

The parameter adjustments Δf_0 and δ are then given by

$$\Delta f_0 = \frac{(\sum w_i \Delta y_i)(\sum w_i \Delta f_i^2) - (\sum w_i \Delta y_i \Delta f_i)(\sum w_i \Delta f_i)}{(\sum w_i)(\sum w_i \Delta f_i^2) - (\sum w_i \Delta f_i)^2} \tag{17}$$

$$\delta = \frac{(\sum w_i \Delta y_i \Delta f_i)(\sum w_i) - (\sum w_i \Delta y_i)(\sum w_i \Delta f_i)}{(\sum w_i)(\sum w_i \Delta f_i^2) - (\sum w_i \Delta f_i)^2}$$

where $\Delta y_i = y_i - f_0 - f(x_i, \beta_1)$.

The weights w_i appearing in equations (15) and (17) should be inversely proportional to the *a priori* variances of the experimental values y_i. We take the variances to be inversely proportional to the number of animals in each experimental group. This ignores the effect of radiation in changing the dispersion of life spans, but a detailed examination shows it to be a not unreasonable procedure.

The theoretical curve obtained in the above manner is shown in Figure 1, with

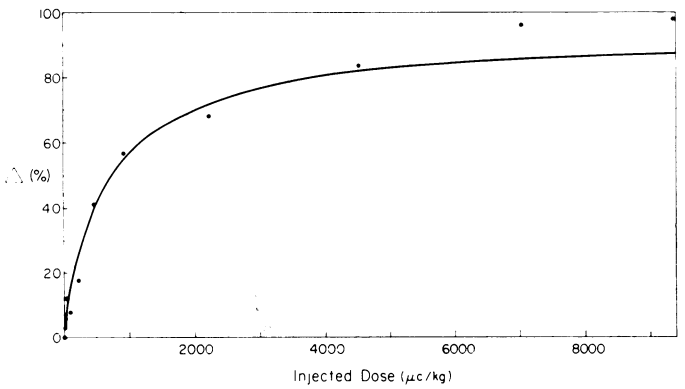

FIG. 1.—Percentage decrease in life expectancy, Δ, as a function of injected dose α^* of strontium-90. Solid curve is the theoretical curve calculated from equation (12). Solid circles are the experimental values reported by Dr. Finkel.

the experimental points for comparison. The two points at highest radiation levels lie well above the curve, doubtless because the mechanism of life shortening at the high radiation levels departs from what we have assumed, owing to the importance of subacute and acute irradiation disease, which Dr. Finkel reports to be the primary cause of death at injected doses above 2200 μc/kg. These points make little contribution to the least squares parameter adjustment, owing to their low weights, and can be omitted without sensibly changing the result. The parameters obtained are $\beta = 1.8 \times 10^{-7}$ day^{-2} (μc/kg retained)$^{-1}$, and $\Delta_0 = 2.5$ per cent.

The least-squares-fitted curve can be used to estimate the death-rate standard deviation σ: for experimental Δ values of unit weight (taken here to be for the control group and "group 12"), the estimated variance of the experimental Δ values is

$$^2 = \frac{1}{m - 2} \sum_{i=1}^{m} w_i [y_i - \hat{f}_0 - f(x_i, \hat{\beta})]^2 \tag{18}$$

where m is the number of experimental points. Equation (18) takes into account the two-parameter adjustment. If M is the number of animals in groups having unit weight, then

$$\hat{\sigma}^2 = M\hat{\sigma}_\Delta^2 \tag{19}$$

From equations (18) and (19) we find $\hat{\sigma} = 191$ days in case the two highest points mentioned above are omitted ($\hat{\sigma} = 222$ days in case they are included). Comparing the value 191 days with the value $\sigma = 161$ days based on the Gompertz death-rate curve (670-day life span) and the value $\hat{\sigma} = 258$ days inferred from Dr. Finkel's data, we see that the theoretical curve fits the experimental values about as well as would be expected from the Gompertz curve, and somewhat better than would have been expected on the basis of Dr. Finkel's acceptance region.

In Figure 2 there is shown the portion of the theoretical curve for the lower radi-

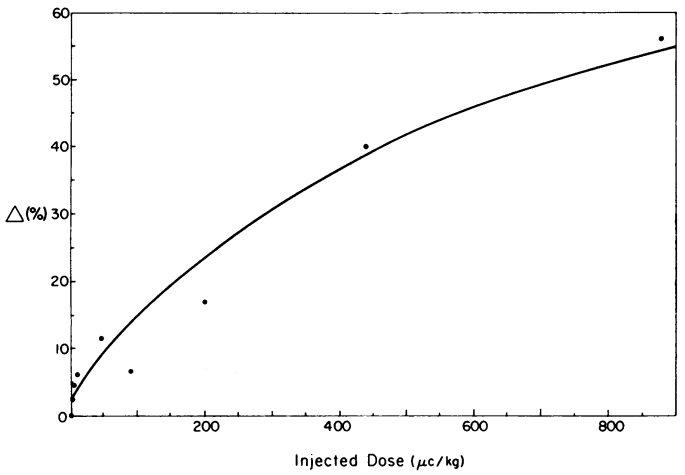

FIG. 2.—Percentage decrease in life expectancy in the low-dose region of Dr. Finkel's experiments. Theoretical curve and experimental points as in Fig. 1.

ation levels. The curvature is pronounced, and the linear response region is restricted to injected doses less than about 50 μc/kg. Most of Dr. Finkel's experiments were carried out in the nonlinear portion of the curve.

It is interesting to compare the above analysis with an alternative one based on the approach developed by Jones,[3] in which the effect of a given exposure of an animal to radiation is regarded as equivalent to an increase in physiological age of the animal by an amount proportional to the amount of radiation received. In terms of the Gompertz formulation of the natural death rate, this results in the case we are considering to the addition of a linear term in t to equation (6):

$$\ln\left(-\frac{1}{N}\frac{dN}{dt}\right) = C + Bt + B\eta\alpha t \tag{20}$$

The constant $B\eta$ in this treatment plays a role analogous to the constant β used previously.

From equations (3) and (20) we obtain

$$t_\alpha = \frac{-\mathrm{Ei}\left(-\dfrac{A}{1+\eta\alpha}\right)}{B(1+\eta\alpha)}\, e^{A/(1+\eta\alpha)} \tag{21}$$

where $A = e^C/B$ as before. In practice A is so small that the exponential integral can adequately be approximated by the logarithmic term in equation (11):

$$-\mathrm{Ei}\left(-\frac{A}{1+\eta\alpha}\right) = \ln\frac{1}{\Gamma A} + \ln(1+\eta\alpha) \tag{22}$$

Recognizing from equations (21) and (22) that

$$t_0 = \frac{1}{B}\ln\frac{1}{\Gamma A}$$

and making use of equation (8), we obtain

$$\Delta = 1 - \frac{1+\epsilon\ln(1+\eta\alpha)}{1+\eta\alpha} \tag{23}$$

where $\epsilon = \tau_D/(t_0 \ln 2) \cong 1/5$.

By choosing $\eta = 0.014$ (μc/kg retained)$^{-1}$, and by adjusting the zero point slightly as done previously, we calculate from equation (23) a theoretical curve that matches closely the curve calculated from equation (12), which is the curve shown in Figures 1 and 2. The discrepancy in Δ between the two curves is less than 0.02 over the range $0 \leq \alpha^* \leq 3000$ μc/kg, and increases to 0.04 at $\alpha^* = 9000$ μc/kg. The two curves fit the experimental data equally well, as shown by the estimates $\sigma = 191$ days for the curve from equation (12) and $\sigma = 189$ days for equation (23), calculated by the weighted sum-of-squares procedure described previously. Life-shortening data, at least of the accuracy involved here, are therefore unable to discriminate between the two analytical approaches.

Analysis of Incidence of Leukemia and Related Diseases.—The experimental data on the incidence of diseases of the blood and blood-forming tissues can be analyzed in the framework of the above treatment. However, because of the peculiar form in which the experimental results are presented ("Curve C: percentage decrease in time to a 20 per cent incidence of reticular tissue tumors compared with the 20 per cent incidence time of the controls"), the analysis is subject to greater uncertainties and difficulties and the data cannot so readily be evaluated statistically as those for the decreased life expectancy. We therefore content ourselves with a somewhat sketchy treatment, which should suffice to indicate the general nature of the problem.

Let $\lambda(t, \alpha)$ be the expected number of deaths due to these diseases that have occurred by the time t in a population having retained body burden α of strontium-90. We may then expect to find a death-rate probability parameter β_l for these diseases such that the death rate is

$$\frac{d\lambda}{dt} = N(t)\beta_l\alpha t + \frac{d\lambda_0}{dt}\frac{N(t)}{N^0(t)}$$

where $\lambda_0(t)$ is the number of deaths due to these diseases expected in the control population. To carry the analysis further we need to know the function $\lambda_0(t)$, but unfortunately Dr. Finkel presents no data that enable us to determine it. Of the various assumptions that could be made, we have chosen to assume that the nat-

ural deaths due to leukemia are distributed as though they were radiation-induced according to the same model as the deaths due to radiation from strontium-90. The natural leukemia death rate will then be equivalent to a "background" body burden α_0 of strontium-90, and equation (20) becomes

$$\frac{d\lambda}{dt} = N(t)\beta_l(\alpha_0 + \alpha)t$$

Obtaining $N(t)$ from equation (2), we have

$$\frac{\lambda(t)}{N_0} = \beta_l(\alpha_0 + \alpha)\int_0^t n^0(t)e^{-(1/2)\beta\alpha t^2}\, t\, dt$$

The expected 20 per cent incidence time τ is then the implicit solution of

$$0.11\beta_l(\alpha_0^* + \alpha^*)\int_0^\tau n^0(t)e^{-(0.11/2)\beta\alpha^* t^2}\, t\, dt = \frac{1}{5} \tag{24}$$

and the expected no-dose 20 per cent incidence time τ_0 is given by

$$0.11\beta_l\alpha_0^*\int_0^{\tau_0} n^0(t)t\, dt = \frac{1}{5} \tag{25}$$

τ_0 as given by equation (25) is not necessarily the same as the 20 per cent incidence time $\tau_0' = 565$ days observed for the control population.

To compare the theory with the experimental data we calculate from equation (24) the fractional decrease function $1 - \tau(\alpha^*)/\tau_0'$. An adequate approximate calculation for values of τ less than about 450 days ($1 - \tau/\tau_0' > 0.20$) can be made by approximating $n^0(t)$ by a Gaussian $e^{-\mu t^2}$, as shown in Table 1. In this case equation (24) becomes

$$0.11\beta_l(\alpha_0^* + \alpha^*) = \frac{\mu + \frac{1}{2}(0.11)\beta\alpha^*}{5\left[1 - \exp\left(-\tau^2\left(\mu + \frac{1}{2}0.11\beta\alpha^*\right)\right)\right]} \tag{26}$$

To evaluate the parameters β_l and α_0^* we have fitted a smooth curve, by eye, to the experimental values of $1 - \tau/\tau_0'$, and used this curve to pick pairs of values (α^*, $\tau(\alpha^*)$) from which the quantity $0.11\beta_l(\alpha_0^* + \alpha^*)$ was calculated from equation (26). The quantity $0.11\ \beta_l\alpha_0^*$ was calculated from equation (25) by numerical integration, with the assumption $\tau_0 = \tau_0'$. When plotted against α^*, the values of $0.11\ \beta_l(\alpha_0^* + \alpha^*)$ calculated in this way lie nicely along a straight line, as required by the theory, for values of α^* in the range $0 \leq \alpha^* \leq 1000$ μc/kg. Above 1000 μc/kg the linear relation breaks down, reflecting the fact that the one experimental value in this higher range, at 2200 μc/kg, lies rather far from the theoretical curve. Ignoring this highest value we obtain in this way the parameters $\beta_l = 0.7 \times 10^{-7}$ day^{-2} (μc/kg)$^{-1}$ and $\alpha_0^* = 200$ μc/kg, from which the theoretical curve shown in Figure 3 is calculated. In addition to the point $\tau(0)$, and the points $\tau(\alpha^*)$ calculated from equation (24) over the range of validity of the Gaussian approximation, we have calculated the slope of the theoretical curve at $\alpha^* = 0$ from the following formula, which can be derived from equation (24):

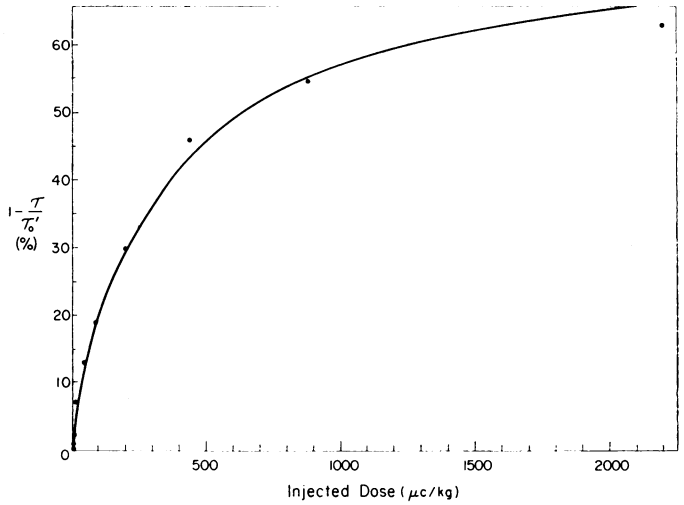

FIG. 3.—Percentage decrease in time to a 20 per cent incidence of bone tumors, as a function of injected dose of strontium-90. Solid curve is calculated from equation (26). Solid circles are experimental values reported by Dr. Finkel.

$$\frac{d}{d\alpha^*}\left(1-\frac{\tau}{\tau_0}\right)_{\alpha^*=0} = \frac{1-\dfrac{\beta}{10\beta_l}\dfrac{\displaystyle\int_0^{\tau_0} n^0(t)t^3dt}{\left[\displaystyle\int_0^{\tau_0} n^0(t)tdt\right]^2}}{10\gamma_0^2\tau_0^2 n^0(\tau_0)}\cdot\frac{1}{2}(0.11)\beta_l$$

where $\gamma_0 = {}^1/_2(0.11)\beta_l\alpha_0{}^*$. The ratio of integrals appearing in the second term of the numerator in this equation can be shown to have a value close to unity (actually 1.06).

A comparison of the parameters $\beta_l = 0.7 \times 10^{-7}$ and $\beta = 1.8 \times 10^{-7}$ suggests that of the radiation-induced deaths the fraction due to leukemia and related diseases in Dr. Finkel's experiments on mice is rather larger than has been estimated for man. A particular sensitivity to these diseases on the part of this strain of mice is suggested also by the large "background dose level" $\alpha_0{}^*$, reflecting the relatively large number of deaths due to these diseases in the control population.

Statistical Examination of Dr. Finkel's Conclusions.—In searching for evidence for the existence of a threshold body burden of strontium-90, below which no harmful effects are caused, Dr. Finkel uses two methods: (1) statistical analysis of the experimental data, and (2) extrapolation of experimental curves. We now consider these two methods.

The statistical analysis consists of a *t*-test of the hypothesis of no difference in response between the control population and a population dosed with strontium-90. Dr. Finkel accepts the null hypothesis at the 10 per cent significance level ("10% probability level or higher") for the three lowest-dosed experimental groups, and considers that this acceptance constitutes "evidence that there might be a threshold" or that "a threshold . . . may lie between 4.5 and 44 μc/kg."

It constitutes nothing of the kind. It is clear that the width of the acceptance

region for the null hypothesis (shaded region in Figs. 3 and 4 of Dr. Finkel's paper) should vary inversely as the square root of the number of animals in the experimental groups, assuming approximate normality of the death-rate curve $g(t)$, as Dr. Finkel must have done in applying the t-test. The threshold for which she finds "evidence" in the experiments is thus no threshold at all but simply a reflection of the statistical uncertainty of her information. It is clear that she could have found "evidence" of this sort for a threshold at any arbitrarily large radiation level (perhaps short of what would produce acute radiation sickness) by simply using few enough animals in her experiments.

The fallacy in Dr. Finkel's statistical argument is a failure to control the probability of type-II error of her test. Type-II error[5] is acceptance of the null hypothesis when it is in fact false. Consideration of the type-II error requires consideration of the alternative to the null hypothesis, which in this case is the theoretically likely linear response at low doses. If we use for the slope $d\Delta/d\alpha^*|_{\alpha^*=0}$ of the life-shortening response at low doses the value obtained above (eq. [13]) from a study of Dr. Finkel's results, namely, $d\Delta/d\alpha^* = 0.14\%$ $(\mu c/kg)^{-1}$, and if we assume that Dr. Finkel's t-test acceptance region is appropriate to a one-tailed test at $\alpha^* = 8.9$ $\mu c/kg$, the highest experimental value for which the null hypothesis was accepted, then we can calculate the probability of type-II error. It is 85 per cent. This means that if there exists in fact no threshold at 8.9 $\mu c/kg$, Dr. Finkel's test would nevertheless have produced "evidence" for one in 85 experiments out of every 100 experiments performed. On the other hand, if there were in fact a threshold, the test would deny it in only 10 per cent of the experiments. Evidently the test is worthless as a proof of the existence of a threshold at this dose level (or lower, for which the probability of type-II error approaches the maximum that is possible, 90 per cent, for a 10 per cent probability of type-I error).

It is incumbent upon those who would extrapolate their threshold conclusions from 150 mice to 3×10^9 human beings that they demonstrate the existence of a significant experimental departure from the theoretically likely linear response, because although the existing burdens of strontium-90 are low, the number of individuals involved is very large, and the harmful consequences of proceeding on an unfounded assumption of a threshold are great.[6] As we have shown above, Dr. Finkel's results are in complete harmony with a linear law; in fact, the agreement between the linear law and the experimental results is better than could have been expected on the basis of the width of her null-hypothesis acceptance region.

As an alternative to the statistical tests, Dr. Finkel determines a threshold by extrapolating the experimental life-shortening curve. She states: "Since the [life-shortening] values for 1.3, 4.5, and 8.9 $\mu c/kg$ do lie along a straight line when plotted semilogarithmically, it may be argued that they represent true departures from the control value. An extension of this straight line crosses the control value at 0.4 $\mu c/kg$." It is difficult to see why the semilog plot rather than some other should be used for the extrapolation. But in fact an extrapolation of any kind is groundless. The three response values lie within less than half the range of probable error (within $-1/2$ P.E. to $+1/2$ P.E.) of the difference d between experimental values of Δ, as determined from the width of the null-hypothesis acceptance region ($\sigma_d = 5.5$ per cent). If it is not obvious that no non-zero regression slope determined from these points can have any statistical significance, one can show[7] that

the standard deviation of the regression slope estimator derived from the semilogarithmic plot is 2.2 times the estimated slope itself. If the semilogarithmically linear relation of the three points can be ascribed to anything but chance, then all of Dr. Finkel's statistical arguments are false. It is hard to imagine how two such mutually contradictory "proofs" could be advanced at one time.

There are other statistical points in Dr. Finkel's paper that merit scrutiny. In our discussion of her results we have had to rely on the correctness of the null-hypothesis acceptance region that she presents, but there are serious reasons for doubting its correctness. The width of the acceptance region corresponds to the estimate $\hat{\sigma} = 284$ days for the standard deviation of the death-rate function $g(t)$, if it applies to a one-tailed test on the difference between the life-shortening values obtained from two experimental groups of 150 animals each. On the other hand, Dr. Finkel's statement[8] that groups of 1393 animals would have been required to establish as significant at the 1 per cent level the difference observed (2.5 per cent) at the lowest dose corresponds to $\hat{\sigma} = 171$ days, a gross discrepancy. The latter value, we note, agrees reasonably with the values $\sigma = 161$ days from the Gompertz relation or $\hat{\sigma} = 191$ days from the agreement between experimental data and our theoretical curve.

It is clear that since the number of animals differs from one experimental group to another in Dr. Finkel's experiments, the null-hypothesis acceptance region cannot have width independent of injected dose α^*, as shown in her figures. From the information given there is no way to tell to which experimental groups the test appropriately applies.

More serious is the evident fact that Dr. Finkel applies the same acceptance region indiscriminately to the three very different sets of experimental data represented by her curves A, B, and C. It seems likely that the test was designed to handle the life-shortening data (curve A), because a t-test would not be inappropriate to life-span data, since the death-rate function $g(t)$ is (rather crudely) Gaussian. A statistical analysis of the curve-C data would be difficult, because the experimental statistic τ (20 per cent incidence time) is cumbersome to handle mathematically, as is evident in our discussion. But it is easy to show that Dr. Finkel's acceptance region is entirely inapplicable to the curve-B data ("proportion of animals that survived the latent period of 150 days and then died with osteogenic sarcomas").

The number of bone-cancer deaths in populations of a given size during a given time interval will be Poisson-distributed, if we neglect variations in population size due to deaths during the first 150 days, which is legitimate, as can be seen from Table 1 or from numbers given by Dr. Finkel, which show that the control group still contained close to 150 animals at $t = 150$ days. Whatever the low-dose regression function for curve B, it is clear from Dr. Finkel's Figure 4 that the expected number ξ of bone-cancer deaths is close to 3 for groups of 150 animals not dosed with strontium-90. To find acceptance regions for the null hypothesis of no significant difference in the number of such deaths between the control population and a dosed population of equal size we therefore find the value of the difference δ_P such that the probability of type-I error (one-tailed test) is P:

$$1 - P = e^{-2\xi} \sum_{\delta=-\infty}^{\delta_P} \xi^{\delta} \sum_{n=-\infty}^{+\infty} \frac{\xi^{2n}}{n!(n+\delta)!} \tag{27}$$

The results of a numerical evaluation of equation (27), in case $\xi = 3$, are $P = 0.10$, $\delta_P = 2.6$; $P = 0.01$, $\delta_P = 5$; $P = 0.001$, $\delta_P \cong 8.5$. Using the fact that the observed number of control deaths was 3 (2 per cent of 150), we find that the upper limit of the acceptance region at 10 per cent significance level is 3.7 per cent of 150. From Dr. Finkel's Figure 4 we therefore see that the highest strontium-90 dose that produced a "statistically non-significant increase" in the number of bone-cancer deaths is 8.9 μc/kg, not 200 μc/kg as stated by her. Her acceptance region represents for curve B a test having 0.1 per cent probability of type-I error. Her entire discussion of the statistical significance of the curve-B data is erroneous.

The Proper Testing of Evidence for a Threshold.—From the above discussion it is clear that a valid statistical test of the null hypothesis of "no response" at a given radiation level or a given dose of strontium-90 must use the type-II error as the basic parameter, rather than the type-I error, as is employed in the standard "cook book" tests, which are designed basically for application to the manufacture of goods. Alternatively stated, the null hypothesis that must be tested in the standard way is the hypothesis that the observed response values are in accord with a linear response curve at low doses.

Our analysis of Dr. Finkel's data for mice enables us to estimate reliably the linear decrease in life expectancy for low doses of strontium-90, and it is therefore possible for us to determine how many animals would have to be used in an experiment in which the mean lifetime of a control group is compared with the mean lifetime of a dosed group in order to establish the existence of a threshold at or above the dose used. We may consider two types of test: (A) the "minimal" test, that is, the test that requires as few animals as possible; (B) the "most powerful" test, which minimizes the probability both of type-I and of type-II error.

The null-hypothesis "no response" for test A is to be accepted if the dosed group exhibits no decrease in life expectancy, or an actual increase, when compared to the control group. Clearly this acceptance region makes the test minimal, because the probability of type-I error is 50 per cent, so that the test gives a neutral decision in case a threshold actually exists. If the number of animals used is greater than required for test A, then a decision as to the existence of a threshold will be more often right than wrong, in case that the threshold does actually exist. At the same time we can protect ourselves adequately against the serious alternative possibility by suitably choosing the type-II error.

Since the expected decrease in life expectancy for the dosed group is $\alpha \cdot d\Delta/d\alpha \big|_0$, the type-I and type-II errors are simultaneously minimized, and made equal, by choosing as the upper limit of the acceptance region for test B a decrease in life expectancy of $^1/_2\alpha \cdot d\Delta/d\alpha \big|_0 = 0.63\% \cdot \alpha$, where α is given in μc/kg retained in the body.

Since the expected decrease is proportional to α, the number of animals M required for the control group, if an equal number is used for the dosed group, is given by

$$M = \frac{\nu}{\alpha^2} \tag{28}$$

where ν is a constant that depends on the type of test (A or B), on the probability of type-II error, and on the standard deviation σ of the natural death-rate function

$g(t)$. Because M proves in all cases of interest to be large, it is adequate to use the normal distribution in computing the constants ν in equation (28), owing to the Central Limit Theorem.

TABLE 2

VALUES OF THE CONSTANT ν IN EQUATION (28)

	TYPE OF TEST			
	A ("Minimal")		B ("Most Powerful")	
σ (days)	284	170	284	170
10 per cent probability of type-II error	4560	1630	18240	5520
1 per cent probability of type-II error	14950	5360	59800	21440

Using all of these principles, we have computed the coefficients ν for the various circumstances shown in Table 2. In particular we compare the results for σ = 284 days, derived from Dr. Finkel's acceptance region, with the results for σ = 170 days, which seems most reasonable on the basis of the previous discussion.

The numbers ν given in Table 2 are equal to the number of animals in the control and in the dosed groups required to establish the existence of a threshold at α^* = 9.1 μc/kg, just above the highest injected dose (8.9 μc/kg) for which Dr. Finkel accepted the hypothesis that a threshold exists. The numbers of animals used in her experiments were too small by factors of 10 to 400, for the conclusion that she reached. By solving equation (28) for α we may compute very simply the lowest threshold α_T^* that could have been recognized with statistical significance in her experiments, assuming that 150 animals were used both in the control group and in the dosed groups, which in general was not the case (fewer were used). These values of α_T^* are 91 and 54 for test A and 181 and 109 for test B (in each case for σ = 284 days and 170 days, respectively). It is clear from the experimental data that no threshold exists at any of these levels, and accordingly we are required to conclude that Dr. Finkel's data show that there is no threshold large enough to have been recognized with statistical significance from her data.

Conclusions about Effects on Man of Strontium-90 from Fallout.—We now turn to the discussion of Dr. Finkel's conclusion that the present contamination with strontium-90 from fallout is so very much lower than the "threshold" levels that it is extremely unlikely to induce even one bone tumor or one case of leukemia.

This statement by Dr. Finkel is shown by the argument given above to have no justification whatever from her experimental results, obtained with 150 mice or fewer in her control group and injected groups. We may ask how many mice would be needed in each group in order to permit Dr. Finkel's statement to be made with statistical significance (or to be shown to be false).

The present average body burden of strontium-90 in the world's population is about 0.0002 μc. per person. This corresponds, with Dr. Finkel's conversion factor (5 to 10 μc. per 70-kg man equivalent to 1 μc. retained per kg for mice) to a retained dose α = 0.00002 to 0.00004 μc/kg in mice. Hence in order to justify Dr. Finkel's statement evidence would be needed that the mouse threshold is as great as about 0.00004 μc/kg; that is, we must place α in equation (28) equal to 0.00004 μc/kg. From the values of the constant ν in Table 2 (we use the values for σ = 170 days, which we believe to be better than those for σ = 284 days) we find $M = 1 \times 10^{12}$ for the "minimal" test and 3.4×10^{12} for the "most powerful" test with 10 per cent type-II error, and 3.3×10^{12} and 13.5×10^{12}, respectively, with 1 per cent type-II

error. We hence conclude that a study like that made by Dr. Finkel would have to use a much greater number of mice than the number of people in the world, in order to provide evidence that would justify her extreme statement to be made with statistical significance.

This conclusion is, of course, not at all unexpected. The difficulty of detecting by statistical methods an effect that causes a small increase in the annual number of deaths among the world's population is well known. For example, let us assume that the average number of deaths per year is 50 million. The statistical fluctuations from this average from year to year are measured roughly by the square root of this number, 7000; and accordingly a study of a larger population would be needed to show with statistical significance the existence of an effect resulting in an additional 1000 deaths per year (the rough estimate of the world-wide effect of the present body burden of strontium-90 from fallout, if there is no threshold). The same number of mice would be needed to test the equivalent effect in mice.

Summary.—We have developed methods of theoretical analysis of the results of experimental studies of the effects of injection of radioactive substances into animals on their life expectancy and on the incidence of tumors. These methods have been applied to the data reported for mice by Dr. Miriam P. Finkel, and it has been shown that her conclusion from these data that it is extremely unlikely that the strontium-90 from the fallout from nuclear weapons tests will induce even one bone tumor or one case of leukemia in human beings is completely unjustified.

* This paper is a contribution from the Division of the Geological Sciences (No. 908) and the Division of Chemistry and Chemical Engineering (No. 2421) of the Institute. A brief account of the work has been published by us (Letter to the Editor, *The New York Times*, Nov. 16, 1958), and a reply has been made by Dr. Finkel (*ibid.*, Nov., 30 1958).

[1] M. P. Finkel, *Science*, **128**, 637, 1958.

[2] E. B. Lewis, *Science*, **125**, 965, 1957.

[3] H. Jones, *The Nature of Radioactive Fallout and Its Effects on Man* (Washington, D. C.: Government Printing Office, 1957), p. 1109.

[4] Jahnke and Emde, *Tables of Functions* (New York: Dover Publications, 1945), pp. 1–9.

[5] A. M. Mood, *Introduction to the Theory of Statistics* (New York: McGraw-Hill Book Co. Inc., 1950), p. 247.

[6] L. Pauling, *No More War!* (Dodd-Mead & Co., 1958).

[7] Mood, *op. cit.*, pp. 292–294.

[8] Finkel, *op. cit.*, pp. 638–639.

---

*GRAVITATIONAL PROPERTIES OF ANTIMATTER**

BY L. I. SCHIFF

DEPARTMENT OF PHYSICS AND INSTITUTE OF THEORETICAL PHYSICS,
STANFORD UNIVERSITY, STANFORD, CALIFORNIA

Read before The Academy, November 7, 1958

Eötvös Experiments and Equivalence Principle.—A very precise series of experiments involving gravity was performed by Eötvös and collaborators between 1890 and 1922.[1] These experiments sought for possible variations in the ratio of gravitational to inertial mass from one substance to another. By "gravitational mass"