tion of mechanisms increasing message translatability is higher (21). The difference between CV1 and HeLa cells suggests that an additional intracellular control mechanism for HTLV-III/LAV replication could be the extent of trans-activation in different cellular environments.

REFERENCES AND NOTES
15. T. S. Roberts et al., in preparation.
29. B.K.F. was supported by a Swiss National Science Foundation grant. H.F. is a fellow of the Leukemia Society of America. We thank A. Aldovini and C. Debuck for the rabbit antibody to trans-activator and F. Wong-Staal and R. Gallo for the HTLV-III/LAV clones and for sharing unpublished data. We also thank G. Kaplin and T. Tse for expert technical assistance, H. Marusidu for typing and editorial assistance, and K. Rivas for editorial assistance. Supported by the National Cancer Institute, under contract N01-CO-23909 with Biosenetics Research, Inc.

3 July 1986; accepted 17 September 1986

Distribution of Airborne Radon-222 Concentrations in U.S. Homes

A. V. Nero, M. B. Schwehr,* W. W. Nazaroff,† K. L. Revzan

Apparently large exposures of the general public to the radioactive decay products of radon-222 present in indoor air have led to systematic appraisal of monitoring data from U.S. single-family homes; several ways of aggregating data were used that take into account differences in sample selection and season of measurements. The resulting distribution of annual-average radon-222 concentrations can be characterized by an arithmetic mean of 1.5 picocuries per liter (55 becquerels per cubic meter) and a long tail with 1 to 3% of homes exceeding 8 picocuries per liter, or by a geometric mean of 0.9 picocuries per liter and a geometric standard deviation of about 2.8. The standard deviation in the means is 15%, estimated from the number and variability of the available data sets, but the total uncertainty is larger because these data may not be representative. Available dose-response data suggest that an average of 1.5 picocuries per liter contributes about 0.3% lifetime risk of lung cancer and that, in the million homes with the highest concentrations, where annual exposures approximate or exceed those received by underground uranium miners, long-term occupants suffer an added lifetime risk of at least 2%, reaching extraordinary values at the highest concentrations observed.

Radon-222 and its decay products are universally present in the air we breathe, with typical levels indoors estimated—on the basis of epidemiology among uranium miners—to cause thousands of cases of lung cancer annually in the United States. A special concern has been the frequent appearance of homes with concentrations that imply individual lifetime risks of lung cancer exceeding 1%, and the occasional occurrence of levels with estimated risks an order of magnitude or more higher. These risks, large compared with ordinarily considered environmental risks, have led to diverse studies characterizing indoor concentrations, the factors affecting them, and health implications (1, 2). However, there has not been a quantitative characterization of the distribution of radon concentrations in U.S. residences. Monitoring has been limited to modest local efforts varying markedly in scientific objectives, selection of homes, and measurement techniques, providing no direct estimate either of average exposures or of the number of homes above proposed action levels, such as the criterion recently recommended by the U.S. National Council on Radiation Protection and Measurements (NCRP): 2 working-months (WLM) per year (3, 4). Nonetheless, taken together, data from U.S. studies are substantial. We present results of a systematic appraisal of these data, designed to aggregate them in a consistent way, explicitly accounting for their differences, and thereby to estimate the frequency distribution of concentrations in U.S. homes. We demonstrate the utility of lognormal parameters for representing this distribution and extract quantitative values for average concentrations and for the incidence of high levels; both lognormal and nonparametric analyses give similar results.

From the literature and from direct communication with researchers, we have accumulated data from 38 U.S. areas, typically urban centers or states (Table 1). Approximately 99% of these data were taken in single-family houses, typically selected by asking for volunteers (for example, from among employees of a given institution) or by choosing from participants in energy conservation programs. Thus, while few, if any, of the homes monitored were selected by statistically based sampling procedures, virtually all selection processes contained a strong random element. The studies varied substantially in size: eight have 50 or more homes and an equal number have fewer than 10. Each data set still gives a useful indication of concentrations in the corresponding area and, taken together with the other data sets, of concentrations in U.S. single-family houses.

Indoor Environment Program, Lawrence Berkeley Laboratory, University of California, Berkeley, CA 94720.

*Present address: California Air Resources Board, Sacramento, CA 95816
†Present address: Department of Engineering Sciences, California Institute of Technology, Pasadena, CA 91125

This content downloaded from 134.88.3.41 on Sun, 03 Apr 2016 17:31:05 UTC
All use subject to http://about.jstor.org/terms
Two broad incentives give rise to most of these data: the first 22 sets came primarily from studies ascertaining present concentrations as a basis for estimating potential increases due to energy-conserving measures that reduce air infiltration into houses. In contrast, sets 23 to 38 were collected in areas with some prior indication of the probability of high $^{222}\text{Rn}$ concentrations. For none of the 38 sets is it thought that use of $^{222}\text{Rn}$-bearing industrial residues or incorporation of unusually effective infiltration-reduction measures have contributed significantly to $^{222}\text{Rn}$ concentrations in the homes monitored.

Nonparametric analysis of these data yields significant information, but because monitoring in large housing samples consistently yields concentrations that are distributed approximately lognormally, we use lognormal parameterization for a more complete and powerful representation of the data. The utility of this functional form is indicated in individual data sets in Fig. 1 and in the distribution obtained by directly aggregating the data from 19 of sets 1 through 22 (Figs. 1 and 2). The $\chi^2$ for the data in Fig. 2 with respect to the indicated lognormal function is large enough so that it or a larger value would occur with only about 5% probability if the function represents the true distribution. This relatively large $\chi^2$ appears to arise from the presence of slightly fewer houses at midrange concentrations (around 3 pCi/liter) than suggested by a lognormal distribution and from the corresponding appearance of more houses in the long tail and with low concentrations. Other functional forms do not provide a good representation of the data.

For each data set, Table 1 states the arithmetic mean (AM), the geometric mean (GM), and the geometric standard deviation (GSD). Study results are expressed in terms of $^{222}\text{Rn}$ concentration, even for studies

<table>
<thead>
<tr>
<th>Location</th>
<th>Houses</th>
<th>$^{222}\text{Rn}$ (pCi/liter)</th>
<th>Measurement protocol</th>
<th>Aggregation code</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>AM</td>
<td>GM</td>
<td>GSD</td>
<td>Technique†</td>
</tr>
<tr>
<td>Washington</td>
<td>138</td>
<td>1.11</td>
<td>0.71</td>
<td>2.48</td>
<td>TD</td>
</tr>
<tr>
<td>Oregon</td>
<td>90</td>
<td>1.27</td>
<td>0.86</td>
<td>2.27</td>
<td>TD</td>
</tr>
<tr>
<td>Montana</td>
<td>20</td>
<td>1.87</td>
<td>1.32</td>
<td>2.34</td>
<td>TD</td>
</tr>
<tr>
<td>Idaho</td>
<td>17</td>
<td>3.56</td>
<td>1.90</td>
<td>2.94</td>
<td>TD</td>
</tr>
<tr>
<td>San Francisco, CA</td>
<td>29</td>
<td>0.45</td>
<td>0.25</td>
<td>2.19</td>
<td>GS</td>
</tr>
<tr>
<td>Las Vegas, NV</td>
<td>3</td>
<td>1.06</td>
<td>0.79</td>
<td>2.47</td>
<td>TD</td>
</tr>
<tr>
<td>Colorado Springs, CO</td>
<td>16</td>
<td>2.54</td>
<td>1.87</td>
<td>2.34</td>
<td>TD</td>
</tr>
<tr>
<td>Fargo, ND</td>
<td>11</td>
<td>7.62</td>
<td>5.69</td>
<td>2.39</td>
<td>TD</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>50</td>
<td>1.45</td>
<td>1.10</td>
<td>2.17</td>
<td>TD</td>
</tr>
<tr>
<td>Houston, TX</td>
<td>103</td>
<td>0.59</td>
<td>0.47</td>
<td>1.94</td>
<td>TD</td>
</tr>
<tr>
<td>Portland, ME</td>
<td>11</td>
<td>0.70</td>
<td>0.44</td>
<td>2.98</td>
<td>TD</td>
</tr>
<tr>
<td>New York</td>
<td>27</td>
<td>1.57</td>
<td>0.89</td>
<td>2.65</td>
<td>TD</td>
</tr>
<tr>
<td>Rochester, NY</td>
<td>8</td>
<td>0.93</td>
<td>0.50</td>
<td>4.08</td>
<td>CR</td>
</tr>
<tr>
<td>Northeastern NY</td>
<td>9</td>
<td>1.33</td>
<td>0.78</td>
<td>3.12</td>
<td>TD</td>
</tr>
<tr>
<td>New York City area</td>
<td>11</td>
<td>0.49</td>
<td>0.45</td>
<td>1.56</td>
<td>PM</td>
</tr>
<tr>
<td>New Jersey</td>
<td>15</td>
<td>1.55</td>
<td>1.36</td>
<td>1.73</td>
<td>PM</td>
</tr>
<tr>
<td>Princeton, NJ</td>
<td>11</td>
<td>2.47</td>
<td>1.09</td>
<td>3.95</td>
<td>TD,GS</td>
</tr>
<tr>
<td>Pittsburgh, PA</td>
<td>122</td>
<td>2.37</td>
<td>1.45</td>
<td>2.35</td>
<td>TD</td>
</tr>
<tr>
<td>Philadelphia, PA area</td>
<td>31</td>
<td>2.64</td>
<td>1.84</td>
<td>2.32</td>
<td>AC</td>
</tr>
<tr>
<td>Damascus, MD</td>
<td>41</td>
<td>3.91</td>
<td>3.04</td>
<td>2.11</td>
<td>AC</td>
</tr>
<tr>
<td>East Tennessee¹</td>
<td>40</td>
<td>3.04</td>
<td>2.0</td>
<td>2.5</td>
<td>TD</td>
</tr>
<tr>
<td>Charleston, SC</td>
<td>20</td>
<td>0.58</td>
<td>0.47</td>
<td>2.12</td>
<td>TD</td>
</tr>
<tr>
<td>Butte, MT²</td>
<td>179</td>
<td>6.39</td>
<td>3.46</td>
<td>3.03</td>
<td>RP</td>
</tr>
<tr>
<td>Grand Junction, CO</td>
<td>104</td>
<td>1.57</td>
<td>1.39</td>
<td>1.61</td>
<td>RP</td>
</tr>
<tr>
<td>Sante Fe, NM</td>
<td>6</td>
<td>1.23</td>
<td>0.94</td>
<td>2.26</td>
<td>RP</td>
</tr>
<tr>
<td>Farmington, NM</td>
<td>6</td>
<td>0.83</td>
<td>0.74</td>
<td>1.66</td>
<td>RP</td>
</tr>
<tr>
<td>New Mexico</td>
<td>9</td>
<td>3.29</td>
<td>2.31</td>
<td>2.45</td>
<td>GS</td>
</tr>
<tr>
<td>Chicago, IL</td>
<td>44</td>
<td>1.02</td>
<td>0.5</td>
<td>3.3</td>
<td>GS</td>
</tr>
<tr>
<td>Central Maine</td>
<td>70</td>
<td>2.97</td>
<td>1.66</td>
<td>2.56</td>
<td>TD</td>
</tr>
<tr>
<td>Lewiston, NY</td>
<td>10</td>
<td>0.52</td>
<td>0.50</td>
<td>1.31</td>
<td>PM</td>
</tr>
<tr>
<td>Monroe, NJ</td>
<td>15</td>
<td>0.54</td>
<td>0.52</td>
<td>1.39</td>
<td>PM</td>
</tr>
<tr>
<td>Canonsburg, PA</td>
<td>8</td>
<td>1.30</td>
<td>0.94</td>
<td>2.15</td>
<td>PM</td>
</tr>
<tr>
<td>Eastern Pennsylvania</td>
<td>36</td>
<td>5.76</td>
<td>3.29</td>
<td>3.41</td>
<td>TD</td>
</tr>
<tr>
<td>Maryland</td>
<td>58</td>
<td>3.80</td>
<td>1.90</td>
<td>3.69</td>
<td>GS</td>
</tr>
<tr>
<td>Oak Ridge, TN</td>
<td>14</td>
<td>1.69</td>
<td>1.28</td>
<td>2.24</td>
<td>GS</td>
</tr>
<tr>
<td>Raleigh, NC</td>
<td>10</td>
<td>0.80</td>
<td>0.52</td>
<td>2.87</td>
<td>GS</td>
</tr>
<tr>
<td>Florida (non-</td>
<td>29</td>
<td>0.81</td>
<td>0.73</td>
<td>1.60</td>
<td>RP</td>
</tr>
<tr>
<td>mineralized)</td>
<td>4</td>
<td>2.75</td>
<td>2.14</td>
<td>2.51</td>
<td>RP</td>
</tr>
</tbody>
</table>

*Measurement technique abbreviations: TD, track-etched detector; GS, grab sample; CR, continuous radon monitor; PM, passive environmental radon monitor; AC, activated carbon integrating device; and RP, radon progeny integrating sampling unit. †Period abbreviations y, m, w, or d represent the continuous number of years, months, weeks, or days of measurement; instantaneous, m, w, or d, measurement performed each season; and 1/m, 1/w, or d, measurement performed each month. ‡Seasons: W, winter; Sp, spring; S, summer; F, fall; and Yr, spans four seasons. $G$ gives normalization and for (sets 1 through 22) population weights. Seasonal data normalization: A, parameters already reflect an estimate of annual average concentration, usually $0.5 \times (\text{winter value} \times \text{summer value})$ where such data exist, aggregated (where needed) with a normalized subset of single-season values; H, heating season results require normalization to obtain annual average. Decimal number gives population weight, approximately equal to population (in millions) within a 50-mile radius centered on a named city or within a state (less any population assigned to cities). No data available on individual houses; obtained only statistical data and for sets 10, 18, 21 the values exceeding 8 pCi/liter (from either graphical information or personal communication). ‡$^{222}\text{Rn}$ decay-product concentration was measured (5).
where decay-product concentration was measured (5). The average (AM) concentration ranges from 0.45 (San Francisco) to 7.6 pCi/liter (Fargo and Eastern Pennsylvania).

Because measurements were performed by a variety of techniques, there may be questions about consistency in overall calibration; and statistical uncertainties associated with individual measurements may be large for the lower concentrations observed. Results of interlaboratory calibration efforts of the last several years (6) indicate that potential inconsistencies in calibration are likely to be small compared with the differences in means (Table 1). The effect of statistical uncertainties on our aggregate parameters is noted below.

Table 1 does not include all data sets accumulated in the United States. For a few we were not able to acquire the data; several sets appear unsuitable for present purposes, and we may be unaware of some data. Moreover, some studies are ambiguously reported with data represented piecemeal or in contradictory fashion; for these we have relied on what appears to be the primary reference. In addition, a number of important studies are still in progress. Nonetheless, Table 1 is a good approximation to a complete representation of available results on U.S. single-family residences.

Aggregating these data sets to approximate a U.S. distribution requires careful attention both to monitoring protocols and to the selection and weighting of housing samples. For the purpose of exposure (and risk) assessment, an integrated year-long measurement is ideal [and practical with etched-track monitors measuring $^{222}$Rn concentration (7)]. Only a few data sets approach the ideal, but many effectively approximate it by deploying integrating $^{222}$Rn or decay-product monitors during every season. It is the numerous studies that employed integrating samplers for only one or two seasons that require the most careful consideration.

In four studies, $^{222}$Rn measurements have been made in a sample of homes during part of the heating season and again during the summer. Winter concentrations and annual averages (estimated as the average of winter and summer results for each house) are approximately lognormally distributed (Table 2). Taking the ratio of annual-average to winter parameters, we find that results vary modestly among these studies, reflecting statistical uncertainties and differences in environmental conditions or house operation. We take the averages of the AM, GM, and GSD ratios, 0.72, 0.81, and 0.89, respectively, as a reasonable basis for inferring annual-average parameters from heating-season results. In aggregations where winter data are normalized to annual averages, we therefore multiply the winter parameters by the corresponding ratios. More detailed information would permit refining this simple approach, but is unlikely to change significantly the overall U.S. distribution that we infer or the specific results that we calculate.

Considering the factors discussed above, we have aggregated the data in several fashions, in each case representing the aggregate distribution conceptually as a weighted sum of individual distributions, $\Sigma w_i \psi_i$, (where $\psi_i$ is the frequency distribution of the $i^{th}$ data set and $w_i$ its weight). Aggregations are distinguished by whether the annual-winter normalization is applied, by which data sets are included, and by how the individual data sets are weighted. We have considered either all 38 sets together (totaling 1377 houses) or sets 1 to 22 (totaling 817 homes), without prior expectation of high concentrations. We expect the latter grouping to yield results that are more representative of the United States as a whole. The data sets have been weighted, not only by the number of homes monitored, but also with equal weights (independent of number) and, for the 22-set grouping, with population weights. Although weighting by number is appropriate for samples drawn from the same population, equal weights give an indication of the U.S. distribution if each study is considered representative of an area (independently of the number of homes monitored). The population-based weighting we have employed associates with each data set from a city the approximate population within 50 miles (taking measurements to be indicative of concentrations in the local area) and with each set from a state, the state’s population (less any population already associated with cities).

We have calculated the parameters of each aggregate distribution using the AM’s, GM’s, and GSD’s in Table 1, rather than aggregating the data directly, since the specific data from some studies are not available and because this approach simplifies adoption of differing weightings and normalizations. The aggregate results are given in Table 3, with the normalized 22-set grouping having an AM of 1.42 to 1.54 pCi/liter (55 to 57 Bq/m$^3$), depending on weighting.

---

Fig. 1. Plot of observed concentration as a function of the percentage of homes having that concentration or less, for an aggregation of 19 sets (see Fig. 2) and for individual exemplary data sets chosen to span the country: Damascus, Maryland (set 20), Wisconsin (set 9), and Washington (set 1), multiplied by factors of 1.2, 0.5, and 0.33, respectively, to visually separate the sets. With this probability scale, lognormally distributed data should lie along a straight line; the indicated lines correspond to the lognormal parameters given in Fig. 2 (for the aggregation) and Table 1 (for the individual sets).
a GM of 0.85 to 0.89 pCi/liter (31 to 33 Bq/m²), and a GSD of 2.6 to 2.9. Table 3 also
gives the percentage of homes with greater than (or equal to) 8 pCi/liter (300 Bq/m²),
calculated from the GM and GSD for each case and also, for the 22-set groupings only,
obtained by applying the indicated weightings and normalizations to the 24 raw mea-
surement results exceeding 8 pCi/liter. For the
normalized 22-set grouping, the fraction of homes exceeding 8 pCi/liter ranges from
1 to 3%, implying that in approximately a million homes exposure rates exceed 2
WLM/year in the limit of 100% occupancy (5). Aggregations that include all 38 data
sets have larger means (and tails) than do comparable 22-set aggregations, confirming
the expectation that sets 23 to 38 tend to have higher concentrations than the others.
Also as expected, with winter data normalized, means and GSD's are lower than for
the unadjusted data sets.

The 38 sets represent 21 states and the
22-set grouping 17 states, both a substantial
number. Although the smaller grouping
may inadequately represent the U.S. Mid-
west and South, every major area is repre-
sented to some extent in both groupings,
including a notable number of major metro-
politan areas. The general method of house
selection, depending primarily on volunteers
and participants in energy-conservation pro-
grams, probably favors middle-class hous-
ing, but low-income housing is represented
[for example, four sets are derived from a
multicity energy-conservation program (8)].
The most significant restriction of the sam-
ple is that it represents only the approxi-
ately 60 million single-family residences
(9) constituting about 70% of the U.S. hous-
ing stock and a somewhat greater per-
centage of the population. Despite its short-
comings, the available sample appears to be
an approximate representation of such homes, subject primarily to the limitation
that only 22 (or 38) areas are sampled.

Results of our analytical approach are
borne out by direct counting of the data.
The comparison is direct for unnormalized
number-weighted aggregations: for the 22-
set grouping, the lognormal parameters sug-
gest that 2.1% (that is, 17) of the 817
homes exceed 8 pCi/liter, whereas the data
themselves contain 24, distributed among
10 of the 22 data sets (five each in Fargo and
Pittsburgh, and one to three in each of eight
other sets). Thus, aggregate lognormal pa-
rameters might yield a slight underestimate
of the percentage in the tail relative to results
from some more fundamental aggregation
approach. On the other hand, the percent-
age obtained by the direct-count method
with normalization (yielding, for example,
2.5% for equal weighting) might be expected
to overestimate the actual tail for the
distribution of annual average concentra-
tions, since, from Table 2, it appears that,
relative to winter measurements, the annual
average distribution is not only scaled down
but has a smaller GSD. Other uncertainties
aside, the two methods, lognormal estima-
tion and counting, might bound the real
percentage in the tail for normalized aggre-
gations; as seen in Table 3, these two results
are quite close for each 22-set grouping.

We also examined the distribution of the
GM's of the individual data sets, which for
the 22 winter-normalized sets appear to
be lognormally distributed with a GSD of
about 2.0. Adding this GSD in quadrature
to a median GSD of 2.2 from the same data
sets yields an overall GSD of 3.0, consistent
with that found by actual aggregation.
Assuming a GSD of approximately 3,
uncertainty formulas for lognormal distribu-
tions yield that the standard uncertainty in
the mean concentration from a truly represen-
tative survey in, for example, 750 homes,
would be about 4%. However, the uncer-

![Fig. 2. Radon-222 concentrations from direct aggregation of the 552 individual data points in 19 sets (1 to 22, less 10, 18, and 21, for which we do not have the individual data). The smooth curves are the lognormal and Gaussian functions corresponding to the indicated parameters (calculated directly from the data) and a Weibull function with parameters determined by fit to the data. The Gaussian and Weibull functions can be made to fit reasonably well to portions of the data, for example, to the peak or to the tail, but not to the entire distribution.]

<p>| Table 2. Comparison of annual-average* to winter parameters. |
|-----------------|-----------------|-----------------|</p>
<table>
<thead>
<tr>
<th>Location</th>
<th>Houses with winter and summer measurements (n)</th>
<th>Annual-average 222Rn (pCi/liter)</th>
<th>Winter-only 222Rn (pCi/liter)</th>
<th>Ratio of annual-average to winter parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pacific NW</td>
<td>32</td>
<td>3.27 2.18 2.25</td>
<td>4.58 2.97 2.41</td>
<td>0.714 0.734 0.924</td>
</tr>
<tr>
<td>Central Maine</td>
<td>58</td>
<td>2.86 1.61 2.52</td>
<td>3.42 1.66 2.81</td>
<td>0.856 0.970 0.897</td>
</tr>
<tr>
<td>Eastern Pennsylvania</td>
<td>36</td>
<td>7.57 3.29 3.41</td>
<td>11.25 4.40 3.81</td>
<td>0.673 0.748 0.895</td>
</tr>
<tr>
<td>Philadelphia, PA area</td>
<td>23</td>
<td>2.73 2.01 2.15</td>
<td>4.20 2.61 2.63</td>
<td>0.650 0.770 0.817</td>
</tr>
<tr>
<td>Average ratio</td>
<td>0.718 0.806 0.886</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*0.5 × (winter + summer).
tainty in an aggregation of the kind considered here is considerably larger, being dominated by questions of sample selection and monitoring techniques. On the more important issue of sample selection, we note that for 22 individual GM’s distributed with a GSD of 2.0 the aggregate mean has a standard error of 15%, corresponding to an uncertainty of 0.13 out of 0.9 pCi/liter (or 0.2 out of the average of 1.5 pCi/liter). The uncertainty in the fraction above 8 pCi/liter can be estimated directly from the number found (about 20) to be (1/V/20) x 2%, assuming the houses found with more than 8 pCi/liter are randomly selected from the total exceeding this level; however, because of sensitivity to the data in individual sets and to the aggregation approach, expressing the fraction as 1 to 3% is a more complete indication of the uncertainty. Estimating the effect of measurement uncertainty on the aggregate parameters, we believe the means and the fraction above 8 pCi/liter to be virtually unaffected, whereas the GSD may have been increased by 0.1 or less, depending on the uncertainty associated with the individual data.

Thus the aggregate parameters from the normalized 22-set grouping are affected only slightly by choice of aggregation procedure, and the intrinsic uncertainty, that is, that associated with having chosen only 22 areas, can be estimated to be about 15% for the means and a somewhat greater percentage for the fraction above 8 pCi/liter. This analysis therefore yields as overall results: AM = 1.5 ± 0.2 pCi/liter (55 ± 7 Bq/m³) and 1 to 3% of houses above 8 pCi/liter (300 Bq/m³), or GM = 0.9 ± 0.1 pCi/liter (34 ± 4 Bq/m³) and GSD of about 2.8. These results characterize concentrations in U.S. single-family homes only to the extent that the homes selected for the 22-set grouping are representative of the housing stock. Ultimately, these results can only be tested in a statistically designed survey.

The distribution of 222Rn concentrations is determined directly by two parameters: (i) the source strength, the rate at which 222Rn enters the indoor atmosphere from soil (the major U.S. source), domestic water, and building materials (10, 11) and (ii) the ventilation rate, which appears to have less variability than the source strength (8, 10, 12). The fact that these rates are determined by a multiplicative combination of numerous underlying factors, such as radium content, moisture, and permeability of the soil, understructure type and construction practice, local temperatures, and winds, leads to the expectation that the concentration distribution found in either a single area or a national survey will resemble a lognormal function, as will an aggregation of results from a moderately large number of individual surveys (13). For the same reasons, ventilation rates are found to be lognormally distributed [GSD about 2.0 (8, 11)]. And, although a major incentive for studying indoor pollutants has been to assess the effect of infiltration-reduction measures designed to lower energy use (14), typical measures reduce infiltration rates only modestly (10 to 20%), with similar increases in indoor concentrations (15). The first-order effect of a broad program of infiltration reduction is therefore to shift the entire 222Rn distribution slightly higher. (To the extent that any housing class is affected more substantially, it is that with higher infiltration rates, which tend to have lower initial 222Rn concentrations.) Finally, the substantial variability (GSD = 2.0) observed in the GM arises primarily from geographic variability in source strength and is a principal consideration in developing methods of localizing high concentrations (10).

Recent analyses of 222Rn decay-product epidemiology and dosimetry, as they apply to environmental exposures, yield estimates of the individual lifetime risk of lung cancer due to annual exposures of 0.2 WLM (the decay-product exposure from 80% occupancy at indoor 222Rn concentrations of 1 pCi/liter) (5) that cluster around 0.2%, ranging over a factor of 3 in either direction (16). Our result of 1.5 pCi/liter in single-family houses therefore corresponds to a risk of 0.3%, and, assuming a concentration of about 0.5 pCi/liter in multifamily dwellings (17), the average estimated risk is slightly lower. Furthermore, it appears that approximately a million single-family homes may have concentrations exceeding 8 pCi/liter, implying an individual lifetime risk greater than approximately 2% to long-term occupants. The occasional house with concentrations exceeding 50 or even 100 pCi/liter causes truly extraordinary risks. Individual risks of 2% exceed by three orders of magnitude the 10−5 ordinarily considered to warrant action by governmental agencies (18); indeed, the average indoor concentration of 222Rn corresponds to a risk exceeding 100 × 10−5. Such risks indicate the need, not only for substantial characterization efforts, but also for the development of a more complete perspective on environmental exposures (19).

The estimated average individual risk of 0.2 to 0.3% corresponds to about 10,000 annual cases of lung cancer in the U.S. population (depending significantly on the population mix and the period of expression for lung cancers induced). Judging from the integrated exposures indicated in the last column of Table 3, a significant fraction of

Table 3. Radon-222 concentration distributions for various data selections, normalizations, and weights.

<table>
<thead>
<tr>
<th>Aggregation type*</th>
<th>Aggregate parameters (pCi/liter)</th>
<th>Fraction of homes &gt;8 pCi/liter</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AM†</td>
<td>GM†</td>
</tr>
<tr>
<td></td>
<td>Counted and aggregated†</td>
<td>Calculated from GM and GSD‡</td>
</tr>
<tr>
<td>Not normalized</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>1.74</td>
<td>0.97</td>
</tr>
<tr>
<td>Equal</td>
<td>1.96</td>
<td>1.01</td>
</tr>
<tr>
<td>Normalized</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>1.42</td>
<td>0.85</td>
</tr>
<tr>
<td>Equal</td>
<td>1.54</td>
<td>0.88</td>
</tr>
<tr>
<td>Population</td>
<td>1.46</td>
<td>0.89</td>
</tr>
<tr>
<td>38 sets (177 houses)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>2.57</td>
<td>1.22</td>
</tr>
<tr>
<td>Equal</td>
<td>2.11</td>
<td>1.07</td>
</tr>
<tr>
<td>Normalized</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number</td>
<td>2.38</td>
<td>1.13</td>
</tr>
<tr>
<td>Equal</td>
<td>1.86</td>
<td>0.98</td>
</tr>
</tbody>
</table>

*Aggregate distributions labeled “±” include all data sets, while those labeled “±” exclude 23 to 38, with prior expectation of high 222Rn concentrations. Not normalized employ the parameters of Table 1 directly, whereas normalized indicates a normalization of sets acquired during heating season only (labeled “H” in Table 1). Alternative weighting schemes are used for aggregating the data sets: number of homes (number); equal weighting (equal); or population weighting (population) for 22 set grouping only, with populations (in millions) from Table 1. For a nonparametric analysis, AM = N−1∑NiAi, where N = ∑Ni. Similarly, for the fraction >8 pCi/liter, we take F(8) = N−1∑NiFi(8), where Fi(8) is the fraction of homes with at least 8 pCi/liter in set i (after normalization). With a lognormal representation, d = p(GM,GSD), for p a lognormal function, and parameters for the aggregate distribution are given by ln GM = N−1∑Ni ln GMi and (ln GSD)2 = N−1∑Ni [(ln GSDi)2 + (ln GMi)2] − (ln GM)2. AMeuc = GM × exp(0.5[ln GSD]2), which follows for a lognormal distribution. The fraction is obtained by integrating the lognormal function. In parentheses is the fraction of total population exposure associated with these homes. Similar integrations yield: 6 to 15% of homes with ≳ 4 pCi/liter, representing 27 to 54% of total exposure, and 19 to 33% with ≳ 2 pCi/liter, representing 54 to 76% of total exposure.
this incidence (about 15%) may be associated with houses having concentrations exceeding 8 pCi/liter. Thus any program to reduce the exposures of those presently at high risk could both reduce individual risk and noticeably decrease total population dose, more or less offsetting the potential increase (10 to 20%) from reducing infiltration rates in the entire housing stock.

Finally, the significant average risk and the substantial incidence of high individual risks suggests the efficacy of a two-part approach to controlling $^{222}\text{Rn}$ decay-product exposures. One part would ensure that general building practice avoids measured or inferred concentrations that significantly increase average levels of $^{222}\text{Rn}$ indoors; the ventilation part of a current industry standard has this objective and serves effectively as a control on the average concentration (20). The second part would provide an explicit exposure limit, not protective of the individual: this corresponds well with the guidelines recommended by the NCRP and international agencies. The present scientific challenge is to develop a better understanding of the fundamental factors affecting radon source strengths, thereby permitting efficient identification of the geographic areas and homes with high concentrations and providing an effective basis for reducing unacceptable levels.

Note added in proof: Reduction of concentrations exceeding 4 pCi/liter, the limit recently recommended by the Environmental Protection Agency, would affect many more homes—about 7%, estimated from our log-normal representation—and would effect a substantially larger reduction in population dose (20a).

REFERENCES AND NOTES

3. A working level month is the $^{222}\text{Rn}$ decay-product exposure incurred from presence in a concentration of 1 working level for 173 hours (5).
5. Combined $^{222}\text{Rn}$ decay-product concentrations, usually linked to the alpha energy that will ultimately be emitted on decay to $^{218}\text{Po}$, are given as (i) the potential alpha energy concentration (PAEC), traditionally given in working level (WL), where 1 WL = 2.08 $\times 10^{-1} \text{ J/m}^2$, or (ii) the equilibrium-equivalent decay-product concentration (EEDC), the concentration of each decay product that, taken together, would yield the PAEC actually present. An EEDC of 101 pCi/liter yields a PAEC of 1 WL. In our analysis, EEDC is converted to $^{222}\text{Rn}$ concentra- tion assuming a ratio of 0.5 (the median of the 0.1 to 0.9 range observed), implying that 1 pCi of $^{222}\text{Rn}$ per liter supports an EEDC of 0.5 pCi/liter (or a PAEC of 0.005 WL). If experienced 100% of the time, a 2.5 WL exposure per year, so that the NCRP’s limit of 2 WL/year can be associated with 8 pCi of $^{222}\text{Rn}$ per liter. (1 pCi/liter equals 87 Bq/m$^3$, the SI unit.)
15. In some circumstances this may not occur; for example, in houses with crawl spaces, if infiltration-reduction efforts emphasize the understructure, both the source strength and ventilation rate may be reduced (W. W. Nazarov and S. M. Doyle, Health Phys. 46, 266 (1983)).
17. The current concentration of sources of $^{222}\text{Rn}$ suggests lower concentrations in high-rise apartment buildings, which is confirmed in countries where these dominate the housing stock. As an example, a survey of the few U.S. data, a single-floor apartment examined by George and Breslin (28) was found to have 6.26 pCi of $^{222}\text{Rn}$ per liter, at the bottom of the observed concentration range in the area studied and similar to outdoor concentrations [T. H. Gesell, Health Phys. 28, 724 (1975)].
19. For example, would include both naturally occurring indoor exposures and $^{222}\text{Rn}$ decay-product exposures that arise from uranium mill tailings, have become the focus of large control programs, and pose perhaps to not millions, of individuals to suffer exposure greater than the NCRP remedial action limit [Environmental Protection Agency, Final Environmental Impact Statement for Standards for the Control of Byproduct Materials from Uranium Ore Processing (Report EPA 520/1-83-008-1, Environmental Protection Center to four houses reported as 1979, Washington, DC, September 1983); appendix C describes the EPA dose-response analysis].
24. J. J. Quackenbos, personal communication.
25. H. M. Prichard, T. F. Gesell, C. T. Hess, C. V. Weifenbach, P. Nyberg, Environ. Int. 8, 83 (1982); measurements were made during air-conditioning season.
26. I. Nitschke, personal communication, Measurements spanning various seasons (with full predominating) were made during 1979-1980, annual average. Data also exist for 32 energy-efficient homes.
28. R. L. Fleischer, A. Mogro-Campero, L. G. Turner, ibid., p. 407; also includes data for 14 energy- efficient homes.

18 June 1986; accepted 15 September 1986