

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

1. J. M. Coffin *et al.*, *Science* 232, 697 (1986).
2. F. Barré-Sinoussi *et al.*, *ibid.* 220, 868 (1983).
3. S. M. Broder and R. C. Gallo, *N. Engl. J. Med.* 311, 1292 (1984).
4. M. Popovic, M. G. Sarngadharan, E. Read, R. C. Gallo, *Science*, 224, 497 (1984); R. C. Gallo *et al.*, *ibid.*, p. 500; M. G. Sarngadharan *et al.*, *ibid.*, p. 506.
5. L. Ratner *et al.*, *Nature (London)* 313, 277 (1985); S. Wain-Hobson *et al.*, *Cell* 40, 9 (1985); R. Sanchez-Pescador *et al.*, *Science* 227, 484 (1985); M. A. Muesing *et al.*, *Nature (London)* 313, 450 (1985).
6. M. A. Gonda *et al.*, *Science* 227, 173 (1985); P. Sonigo *et al.*, *Cell* 42, 369 (1985).
7. S. K. Arya *et al.*, *Science* 229, 69 (1985); L. Siegel *et al.*, *Virology* 148, 228 (1986).
8. J. Sodroski *et al.*, *Science* 227, 171 (1985); J. Sodroski, R. Patarca, C. Rosen, F. Wong-Staal, W. Haseltine, *ibid.* 229, 74 (1985).
9. A. G. Fisher, E. Collalti, L. Ratner, R. C. Gallo, F. Wong-Staal, *Nature (London)* 316, 262 (1985).
10. B. H. Hahn *et al.*, *ibid.* 312, 166 (1984).
11. F. J. Graham and A. J. Van der Eb, *Virology* 52, 456 (1973).
12. C. M. Gorman, S. L. Moffat, B. H. Howard, *Mol. Cell. Biol.* 2, 1044 (1982); C. M. Gorman, G. T. Merlino, M. C. Willingham, I. Pastan, B. H. Howard, *Proc. Natl. Acad. Sci. U.S.A.* 79, 6777 (1982).
13. A. Aldovini *et al.*, *Proc. Natl. Acad. Sci. U.S.A.* 83, 6672 (1986).
14. W. C. Goh, G. Rosen, J. Sodroski, D. D. Ho, W. A. Haseltine, *J. Virol.* 59, 181 (1986).
15. T. S. Robins *et al.*, in preparation.
16. J. Sodroski *et al.*, *Nature (London)* 321, 412 (1986).
17. C. A. Rosen, J. G. Sodroski, W. A. Haseltine, *Cell* 41, 813 (1985).
18. A. J. Berk and P. A. Sharp, *ibid.* 12, 721 (1977); R. F. Weaver and C. Weissman, *Nucleic Acids Res.* 7, 1175 (1979).
19. B. K. Felber, H. Paskalis, C. Kleinman-Ewing, F. Wong-Staal, G. N. Pavlakis, *Science* 229, 675 (1985).
20. K. A. Jones, J. T. Kadonaga, P. A. Luciw, R. Tjian, *ibid.* 232, 755 (1986).
21. G. A. Papanicolaou *et al.*, *Nature (London)* 319, 555 (1986).
22. B. Cullen, *Cell* 46, 973 (1986).
23. D. Hultmark, R. Klemenz, W. J. Gehring, *ibid.* 44, 429 (1986); E. D. Lewis and J. L. Manley, *Mol. Cell. Biol.* 5, 2433 (1985); S. L. Mansour, T. Ghodzicker, R. Tjian, *ibid.* 6, 2684 (1986).
24. P. J. Southern and P. Berg, *J. Mol. Appl. Genet.* 1, 327 (1982).
25. G. N. Pavlakis and D. H. Hamer, *Proc. Natl. Acad. Sci. U.S.A.* 80, 397 (1983).
26. J. P. Chamberlain, *Anal. Biochem.* 98, 132 (1979).
27. V. Glisin, R. Crkvenjakov, C. Byus, *Biochemistry* 13, 2633 (1974).
28. T. Maniatis, E. F. Fritsch, J. Sambrook, *Molecular Cloning—A Laboratory Manual* (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
29. B.K.F. was supported by a Swiss National Science Foundation grant. H.P. is a fellow of the Leukemia Society of America. We thank A. Aldovini and C. Debouk 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. Marusiodis for typing and editorial assistance, and K. Bivens for editorial assistance. Supported by the National Cancer Institute, under contract NO1-CO-23909 with Bionetics 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 picocurie 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 picocurie 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.

**R**ADON-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-level 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.

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  $^{226}\text{Ra}$ -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 1. Radon-222 and decay-product (5) concentrations measured in U.S. residences. Where individual data are available, primary use is made of main-floor living-room or of averaged living-area values, from which the arithmetic mean (AM), geometric mean (GM), and geometric standard deviation (GSD) are calculated; otherwise these parameters are taken from the original paper or (in a few cases) derived from histograms. (Data sets 23 through 38 had some prior indication of the probability of high concentrations.)

Location	Houses (n)	$^{222}\text{Rn}$ (pCi/liter)			Measurement protocol			Aggregation codes§	Reference
		AM	GM	GSD	Technique*	Period†	Season‡		
1 Washington	138	1.11	0.71	2.48	TD	3m	W	H-4.1	(21)
2 Oregon	90	1.27	0.86	2.27	TD	3m	W	H-2.6	(21)
3 Montana	20	1.87	1.32	2.34	TD	3m	W	H-0.8	(21)
4 Idaho	17	3.56	1.30	2.94	TD	3m	W	H-0.9	(21)
5 San Francisco, CA	29	0.45	0.35	2.19	GS	I	S	4.0	(10)
6 Las Vegas, NV	3	1.06	0.79	2.47	TD	1y	Yr	0.5	(22)
7 Colorado Springs, CO	16	2.54	1.87	2.34	TD	4-5m	WSp	H-0.5	(8)
8 Fargo, ND	11	7.62	5.69	2.39	TD	4-5m	WSp	H-0.2	(8)
9 Wisconsin	50	1.45	1.10	2.17	TD	3-4m	W	H-4.7	(23)
10 Houston, TX	103	0.59	0.47	1.94	TD	3-6m	SF	3.1	(24)
11 Portland, ME	11	0.70	0.44	2.98	TD	4-5m	WSp	H-0.5	(8)
12 New York	27	1.57	0.89	2.65	TD	2-9m	SF <sup>25</sup>	8.5	(25)
13 Rochester, NY	8	0.93	0.50	4.08	CR	2w	W	H-1.0	(26)
14 Northeastern NY	9	1.33	0.78	3.12	TD	~6m	FWSp	H-1.0	(27)
15 New York City area	11	0.49	0.45	1.56	PM	1w/seas	2Yr	7.1	(28)
16 New Jersey	9	1.55	1.36	1.73	PM	1w/seas	2Yr	4.0	(28)
17 Princeton, NJ	11	2.47	1.09	3.95	TD,GS	~5m	WSp	H-1.0	(29)
18 Pittsburgh, PA	122	2.37	1.45	2.35	TD	1y	Yr	2.9	(30)
19 Philadelphia, PA area	31	2.64	1.84	2.32	AC	3d/seas	WS	A-7.1	(31)
20 Damascus, MD	41	3.91	3.04	2.11	AC	3d/seas	WSp	H-5.5	(31)
21 East Tennessee	40	3.04	2.0	2.5	TD	2x3m	WF	H-0.8	(32)
22 Charleston, SC	20	0.58	0.47	2.12	TD	4-5m	WSp	H-0.5	(8)
23 Butte, MT	179	6.39¶	3.46¶	3.03¶	RP	1w/seas	Yr		(33)
24 Grand Junction, CO	62	1.57¶	1.39¶	1.61¶	RP	1w/1m	Yr		(34)
25 Sante Fe, NM	6	1.23¶	0.94¶	2.26¶	RP	1-16d	Yr		(35)
26 Farmington, NM	6	0.83¶	0.74¶	1.66¶	RP	1-12d	SpS		(35)
27 New Mexico	9	3.29¶	2.31	2.45	GS	I	Sp		(12)
28 Chicago, IL	44	1.02	0.5	3.3	GS	I	F		(36)
29 Central Maine	70	2.97	1.66	2.56	TD	6-8m	1-2Yr	A	(37)
30 Lewiston, NY	10	0.52	0.50	1.31	PM	1-2w/seas	2Yr		(38)
31 Middlesex, NJ	15	0.54	0.52	1.39	PM	1-2w/seas	2Yr		(38)
32 Canonsburg, PA	8	1.30	0.94	2.15	PM	1-2w/seas	2Yr		(38)
33 Eastern Pennsylvania	36	7.57	3.29	3.41	TD	2-8m	WS	A	(39)
34 Maryland	58	3.80¶	1.90	3.69	GS	I	SpSF		(40)
35 Oak Ridge, TN	14	1.69¶	1.28¶	2.24¶	GS	I	W	H	(41)
36 Raleigh, NC	10	0.80¶	0.52	2.87	GS	I	F		(42)
37 Florida (non-mineralized)	29	0.81¶	0.73¶	1.60¶	RP	1w/seas	Yr		(43)
38 Florida (mineralized)	4	2.75¶	2.14¶	2.51¶	RP	1w/seas	Yr		(43)

\*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; I, instantaneous; /seas, measurement performed each season; and /1m, each month. ‡Seasons: W, winter; Sp, spring; S, summer; F, fall; and Yr, spans four seasons. §Gives normalization and (for sets 1 through 22) population weights. Seasonal data normalization: A, parameters already reflect an estimate of annual average concentration, usually  $0.5 \times (\text{winter value} + \text{summer value})$  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 impor-

tant 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  $^{222}\text{Rn}$  concentration (7)]. Only a few data sets approach the ideal, but many effectively approximate it by deploying integrating  $^{222}\text{Rn}$  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,  $^{222}\text{Rn}$  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 (Ta-

ble 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,  $(\sum_i w_i d_i) / \sum_i w_i$  (where  $d_i$  is the frequency distribution of the  $i^{\text{th}}$  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 (53 to 57 Bq/m<sup>3</sup>), depending on weighting,

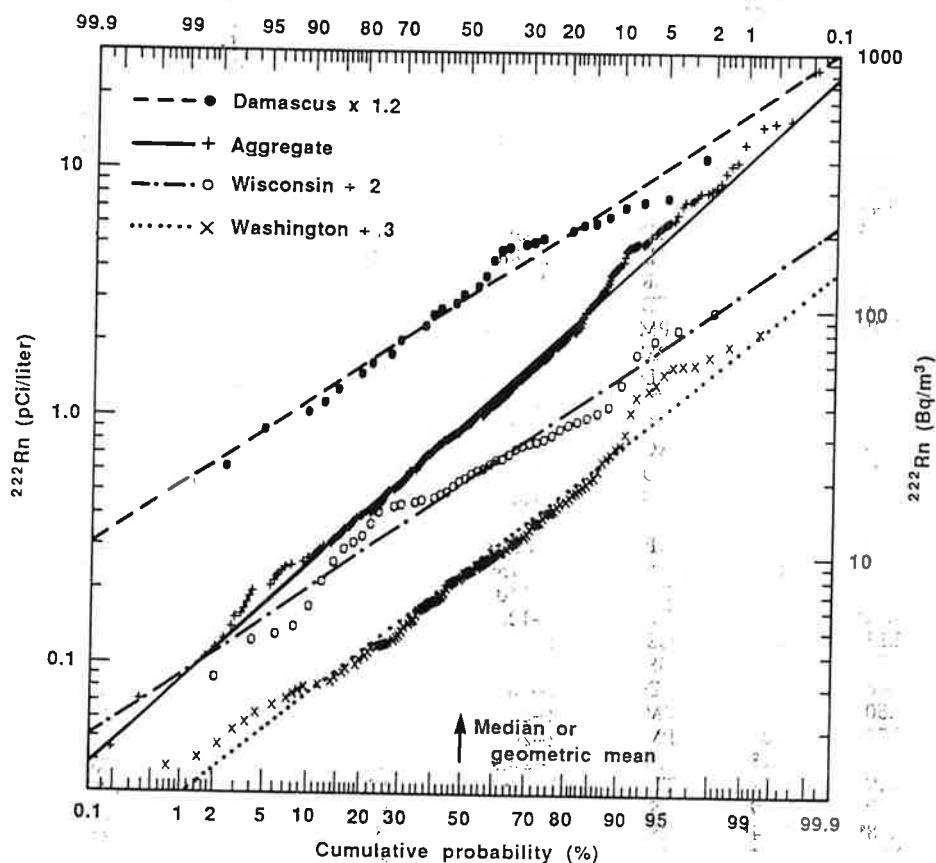


Fig. 1. Plot of observed concentration as a function of the percentage of homes having that concentration or less, for an aggregation of 19 sets (see Fig. 2) and for individual exemplary data sets chosen to span the country: Damascus, Maryland (set 20), Wisconsin (set 9), and Washington (set 1), multiplied 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/m<sup>3</sup>), 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/m<sup>3</sup>), 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 measurement 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. Midwest and South, every major area is represented to some extent in both groupings, including a notable number of major metropolitan areas. The general method of house selection, depending primarily on volunteers and participants in energy-conservation programs, probably favors middle-class housing, 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 sample is that it represents only the approximately 60 million single-family residences (9) constituting about 70% of the U.S. housing stock and a somewhat greater percentage of the population. Despite its shortcomings, 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 supported by specific characteristics of the data. For example, the implication of the lognormal parameters that about 2% of homes have levels exceeding 8 pCi/liter is

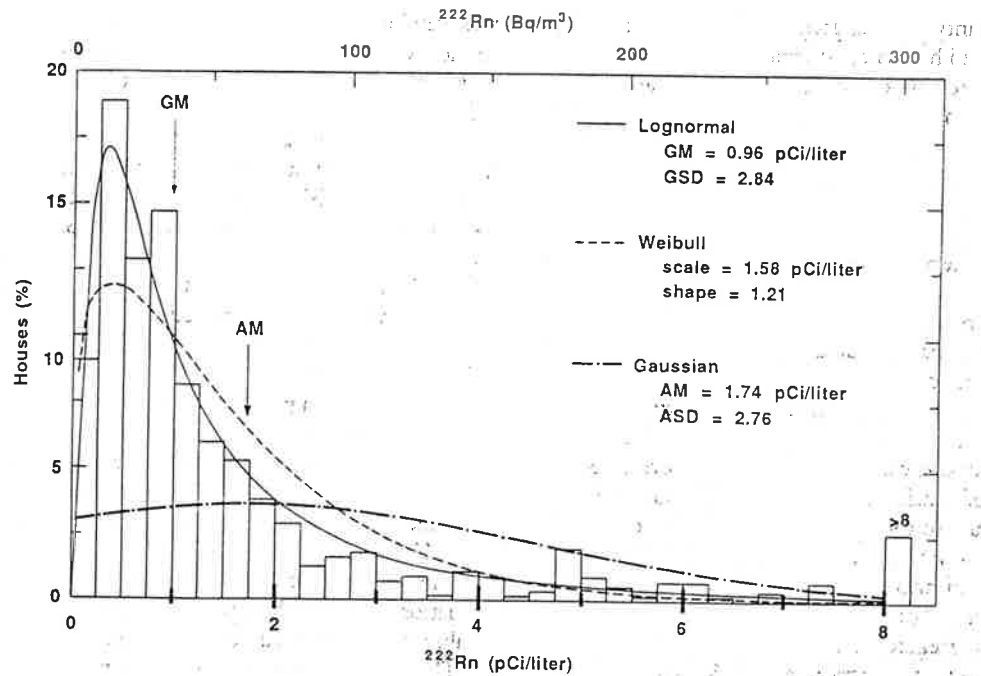


Fig. 2. Radon-222 concentrations from direct aggregation of the 552 individual data points in 19 sets (1 to 22, less 10, 18, and 21, for which we do not have the individual data). The smooth curves are the lognormal and Gaussian functions corresponding to the indicated parameters (calculated directly 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.

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 suggest 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 parameters 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 percentage 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 concentrations, 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 estimation and counting, might bound the real percentage in the tail for normalized aggregations; 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 distributions yield that the standard uncertainty in the mean concentration from a truly representative survey in, for example, 750 homes, would be about 4%. However, the uncer-

Table 2. Comparison of annual-average\* to winter parameters.

Location	Houses with winter and summer measurements (n)	Annual-average <sup>222</sup> Rn (pCi/liter)			Winter-only <sup>222</sup> Rn (pCi/liter)			Ratio of annual-average to winter parameters			Reference
		AM	GM	GSD	AM	GM	GSD	AM	GM	GSD	
Pacific NW	32	3.27	2.18	2.25	4.58	2.97	2.41	0.714	0.734	0.934	(21)
Central Maine	58	2.86	1.61	2.52	3.42	1.66	2.81	0.836	0.970	0.897	(37)
Eastern Pennsylvania	36	7.57	3.29	3.41	11.25	4.40	3.81	0.673	0.748	0.895	(39)
Philadelphia, PA area	23	2.73	2.01	2.15	4.20	2.61	2.63	0.650	0.770	0.817	(31)
						Average ratio		0.718	0.806	0.886	

\*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  $(1/\sqrt{20}) \times 2\%$ , 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 =  $1.5 \pm 0.2$  pCi/liter ( $55 \pm 7$  Bq/m<sup>3</sup>) and 1 to 3% of houses above 8 pCi/liter ( $300$  Bq/m<sup>3</sup>), or GM =  $0.9 \pm 0.1$  pCi/liter ( $34 \pm 4$  Bq/m<sup>3</sup>) and GSD of about 2.8. These results characterize concentrations in U.S. single-family houses 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 <sup>222</sup>Rn concentrations is determined directly by two parameters: (i) the source strength, the rate at which <sup>222</sup>Rn 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 <sup>222</sup>Rn 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 <sup>222</sup>Rn 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 <sup>222</sup>Rn 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 <sup>222</sup>Rn 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<sup>-5</sup> ordinarily considered to warrant action by governmental agencies (18); indeed, the average indoor concentration of <sup>222</sup>Rn corresponds to a risk exceeding 100 × 10<sup>-5</sup>. 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.

Aggregation type*	Aggregate parameters (pCi/liter)				Fraction of homes >8pCi/liter	
	AM†	GM‡	GSD‡	AM <sub>calc</sub> ‡	Counted and aggregated†	Calculated from GM and GSD§
22 sets (817 homes)						
Not normalized						
Number	1.74	0.97	2.81	1.65	0.029	0.021 (0.16)
Equal	1.96	1.01	3.14	1.94	0.045	0.035 (0.25)
Normalized						
Number	1.42	0.85	2.64	1.36	0.018	0.010 (0.09)
Equal	1.54	0.88	2.91	1.56	0.025	0.020 (0.16)
Population	1.46	0.89	2.68	1.45	0.023	0.013 (0.11)
38 sets (1377 homes)						
Not normalized						
Number	2.57	1.22	3.14	2.35		0.051 (0.31)
Equal	2.11	1.07	3.08	2.01		0.037 (0.25)
Normalized						
Number	2.38	1.13	3.09	2.14		0.042 (0.27)
Equal	1.86	0.98	2.96	1.77		0.026 (0.20)

\*Aggregate distributions labeled "38" include all data sets, while those labeled "22" exclude 23 to 38, with prior expectation of high <sup>222</sup>Rn 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). Alternate 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 =  $N^{-1} \sum w_i AM_i$ , where  $N = \sum w_i$ . Similarly, for the fraction  $\geq 8$  pCi/liter, we take  $F(8) = N^{-1} \sum w_i F_i(8)$ , where  $F_i(8)$  is the fraction of homes with at least 8 pCi/liter in set  $i$  (after normalization). With a lognormal representation,  $F_i = p(GM_i, GSD_i)$ , for  $p$  a lognormal function, and parameters for the aggregate distribution are given by  $\ln GM = N^{-1} \sum w_i \ln GM_i$  and  $(\ln GSD)^2 = N^{-1} \sum w_i [(\ln GSD_i)^2 + (\ln GM_i)^2] - (\ln GM)^2$ . †AM<sub>calc</sub> = GM × exp[0.5(ln GSD)<sup>2</sup>], 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  $\geq 4$  pCi/liter, representing 27 to 54% of total exposure, and 19 to 33% with  $\geq 2$  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 measures 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 protecting 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

1. A. V. Nero and W. M. Lowder, Eds., *Health Phys.* 45 (1983), pp. 273–361.
2. G. F. Clemente, H. Eriskat, M. C. O'Riordan, J. Sinnaeve, Eds., *Radiat. Prot. Dosim.* 7, 5 (1984).
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).
4. National Council on Radiation Protection and Measurement, *Exposures from the Uranium Series with Emphasis on Radon and its Daughters* (Report 77, Bethesda, MD, 1984).
5. Combined  $^{222}\text{Rn}$  decay-product concentrations, usually linked to the alpha energy that will ultimately be emitted on decay to  $^{210}\text{Pb}$ , are given as (i) the potential alpha-energy concentration (PAEC), traditionally given in working level (WL), where  $1\text{ WL} = 2.08 \times 10^{-5}\text{ J/m}^3$ , 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  $^{222}\text{Rn}$  concentration 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, 0.005 WL causes 0.25 WLM exposure per year, so that the NCRP's limit of 2 WLM/year can be associated with 8 pCi of  $^{222}\text{Rn}$  per liter. (1 pCi/liter equals  $37\text{ Bq/m}^3$ , the SI unit.)
6. See, for example, I. M. Fisenne, A. George, M. McGahan, *Health Phys.* 45, 555 (1983).
7. H. W. Alter and R. L. Fleischer, *ibid.* 40, 693 (1981).
8. S. M. Doyle, W. W. Nazaroff, A. V. Nero, *ibid.* 47, 579 (1984).
9. Energy Information Administration, *Housing Characteristics, 1982* (Report DOE/EIA-0314 (82), Government Printing Office, Washington, DC, August 1984).
10. A. V. Nero and W. W. Nazaroff, *Radiat. Prot. Dosim.* 7, 23 (1983).
11. W. W. Nazaroff, S. M. Doyle, A. V. Nero, R. G. Sextro, *Health Phys.*, in press.
12. A. V. Nero, M. L. Boegel, C. D. Hollowell, J. G. Ingersoll, W. W. Nazaroff, *ibid.* 45, 401 (1983); set 27 mainly from solar homes.
13. J. Aitchison and J. A. C. Brown, *The Lognormal Distribution* (Cambridge Univ. Press, Cambridge, 1957).
14. A. V. Nero, *Health Phys.* 45, 277 (1983).
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. Nazaroff and S. M. Doyle, *Health Phys.* 48, 265 (1985)].
16. A. C. James, *Radiat. Prot. Dosim.* 7, 353 (1983); N. H. Harley, *ibid.*, p. 371; W. Jacobi, *ibid.*, p. 395; N. H. Harley and W. H. Ellett, in *Indoor Air and Human Health*, R. B. Gammage and S. V. Kaye, Eds. (Lewis, Chelsea, MI, 1985). The highest risk estimate is that of the U.S. Environmental Protection Agency—about 1%. See also (19).
17. The current understanding of sources of  $^{222}\text{Rn}$  suggests lower concentrations in high-rise apartments, which is confirmed in countries where these dominate the housing stock. As an example of the few U.S. data, a single third-floor apartment examined by George and Breslin (28) was found to have 0.26 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.* 45, 289 (1983)].
18. See, for example, C. R. Cothem, W. L. Lappenbusch, J. A. Corruvo, *ibid.* 44 (suppl. no. 1), 377 (1983); see also *Fed. Regist.* 44, 68703 (29 November 1979) for a discussion of risk assessment associated with control of trihalomethanes in drinking water.
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 cause perhaps tens, 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 Agency, Washington, DC, September 1983); appendix C describes the EPA dose-response analysis].
20. American Society of Heating, Refrigerating, and Air-Conditioning Engineers, *Ventilation for Acceptable Indoor Air Quality* (Standard 62-1981, Atlanta, GA, 1981).
- 20a. Environmental Protection Agency, *A Citizen's Guide to Radon* (Report OPA-86-004, Washington, DC, August 1986).
21. P. W. Thor, personal communication.
22. P. C. Nyberg and D. E. Bernhardt, *Health Phys.* 45, 539 (1983).
23. J. J. Quackenboss, personal communication.
24. H. M. Prichard, T. F. Gesell, C. T. Hess, C. V. Weiffenbach, P. Nyberg, *Environ. Int.* 8, 83 (1982); measurements were made during air-conditioning season.
25. I. Nitschke, personal communication. Measurements spanning various seasons (with fall predominating) were taken to approximate an annual average. Data also exist for 32 energy-efficient homes.
26. W. W. Nazaroff, F. J. Offermann, A. W. Robb, *Health Phys.* 45, 525 (1983).
27. R. L. Fleischer, A. Mogro-Campero, L. G. Turner, *ibid.*, p. 407; also includes data for 14 energy-efficient homes.
28. A. C. George and A. J. Breslin, in *Natural Radiation Environments III*, T. F. Gesell and W. M. Lowder, Eds. (CONF-780422, National Technical Information Service, Springfield, VA, 1980), vol. 2, pp. 1272–1292.
29. T. L. Hernandez and J. W. Ring, *Environ. Int.* 8, 45 (1982).
30. B. L. Cohen, *Health Phys.* 49, 1053 (1985).
31. A. C. George, M. Duncan, H. Franklin, *Radiat. Prot. Dosim.* 7, 291 (1984); A. C. George, personal communication. Data set for Philadelphia uses the average of winter and summer concentrations, where available, aggregated with a subset of winter-only concentrations normalized to an annual average (using GM and GSD ratios derived from the rest of the Philadelphia data set).
32. A. R. Hawthorne, R. B. Gammage, C. S. Dudney, in *Indoor Air*, B. Berglund, T. Lindvall, J. Sundell, Eds. (Swedish Council for Building Research, Stockholm, Sweden, 1984), vol. 2, p. 137; A. R. Hawthorne, personal communication. Some measurements in basements are included. Stated GM and GSD are averages of values for two monitoring periods.
33. L. L. Lloyd, personal communication; GM and GSD were calculated from a fit to the number of living-space measurements in four concentration intervals.
34. The GM and GSD were derived by aggregating two data sets of "background" houses from remedial action survey: 29 houses (GM 1.44, GSD 1.7) cited along with cumulative probability plot [B. H. Peterson, in *Radon Workshop, February 1977*, A. J. Breslin, Ed. (Energy Research and Development Administration Report HASL-325, National Technical Information Service, Springfield, VA, 1977), p. 108] and 33 houses (GM 1.35, GSD 1.54 calculated from individual data) (G. A. Franz, personal communication).
35. R. L. Douglas, J. M. Hans, T. A. Wolff, *Working Level Screening Survey of Structures Constructed of Materials Containing Pumice* (Report PB-282-446, Department of Commerce, Washington, DC, 1978); most Santa Fe homes contained pumice block.
36. J. Rundo, F. Markun, N. J. Plondke, J. Y. Sha, M. A. Essling, in *Natural Radiation Environment*, K. G. Vohra, U. C. Mishra, K. C. Pillai, S. Sadasivan, Eds. (Wiley Eastern, New Delhi, 1982), pp. 155–162.
37. C. T. Hess, R. L. Fleischer, L. G. Turner, *Health Phys.* 49, 65 (1985). Data set for central Maine uses average of winter and summer concentrations, where available, aggregated with a subset of summer-only concentrations normalized to an annual average with GM and GSD ratios derived from the rest of the Maine data set. (Half of homes were known to have high levels of  $^{222}\text{Rn}$  in water supplies.)
38. A. C. George and J. Eng, *Health Phys.* 45, 397 (1983); these data are among sets 23 to 38 because of association with potential contamination, but this does not imply contamination of these houses.
39. H. M. Sachs, T. L. Hernandez, J. W. Ring, *Environ. Int.* 8, 97 (1982). We averaged winter and summer values from living areas; we assigned values of 0.1 pCi/liter to four houses reported as zero.
40. D. J. Moschandreas, H. E. Rector, P. O. Tierney, *Geomat. Technol. Rep. GTI ES-877* (1981); D. J. Moschandreas and H. E. Rector, *Environ. Int.* 8, 77 (1982). We omitted two data points due to sampling problems and assigned values of 0.1 pCi/liter to two houses reported as zero.
41. W. M. Lowder, A. C. George, C. V. Gogolak, A. Blay, *U.S. Atom. Energy Comm. Rep. HASL TM 71-8* (1971).
42. L. T. Caruthers and A. W. Waltner, *Health Phys.* 29, 814 (1973). We adjusted the data assuming 0.5 equilibrium factor; activity was underestimated slightly due to improper counting-time analysis.
43. Florida Department of Health and Rehabilitative Services, *Study of Radon Daughter Concentrations in Structures in Polk and Hillsborough Counties* (Tallahassee, 1978). Data were used for houses on non-mineralized ("background") land and on land known to be mineralized (but not mined and reclaimed). [A slightly different set of 29 "background" houses (with GM 0.56 pCi/liter, GSD 1.82) is reported by R. J. Guimond, W. H. Ellett, J. E. Fitzgerald, S. T. Windham, P. A. Cuny, EPA Report 520/4-78-013 (1979).]
44. We are grateful to those who provided us with data before publication. Supported by Department of Energy contract DE-AC03-76SF00098.

18 June 1986; accepted 15 September 1986