

plications for realty transfer. Perhaps the major "market" for radon testing, we see here that realty decisions based upon short-term measurements are likely to suffer from the same extreme level of error as do the measurements themselves.

Radon Testing in Water and Soil

Many of the validity threats seen with testing for radon in air also occur when we examine radon in water and soil. We previously noted the tendency to overweight the significance of radon in water, which is in fact unlikely to be a problem in the absence of an airborne radon problem.¹² Still, radon in groundwater has been recognized as a serious issue (see chapters 3, 4, and 10) given values as high as one million pCi/l. Air tests are frequently used to screen for waterborne radon; a low value suggests that little radon is being released from the water supply. Approaches that directly measure radon in water include laboratory analysis of water samples using a liquid scintillation method or by counting gamma emissions, measuring the outgased radon from a sample of water with an electret ion chamber (Kotrappa and Jester, 1993), and placing a special integrating alpha track detector in the toilet tank for several months (NYSEO, 1989). These tests all suffer from various validity errors. However, of greatest interest with radon tests in water are interpretation errors. At a general conversion of 6,000–10,000 pCi/l in water to 1 pCi/l in air, 24,000–40,000 pCi/l in water would be required before the guidance level of 4 pCi/l in air would be triggered. The state of Maine has utilized 20,000 pCi/l in water as a level suggesting the need for mitigation. Nevertheless, the sheer size of these radon amounts can easily frighten an ill-informed homeowner into carrying out water system mitigations not called for by airborne radon levels (NYSEO, 1989).

Soil testing also confronts serious limitations. While it is simple to test soil for radon using grab samples or with integrating devices placed in the soil for a period of time, these tests have little predictive validity. Most critical are measurement or sampling problems due to the variability of radon levels in soil over very short distances and over time. As a result, the number of tests required to accurately represent a complex soil geology is quite large. Furthermore, even if the site is adequately sampled prior to construction, disturbance of the ground may alter the emanation of radon at a given location and make the transport of radon much easier. Of course, soil radon values are only one factor in determining what a house radon value will be. Soil permeability, weather conditions, and house construction and operation are also important (see chapter 5).

Conclusions about the Myth of the Quick Test

We have seen that, of the six types of error discussed, the validity of radon test-

ing is most affected by fundamental policy and methodological errors leading to interpretation error. Simply stated, the combination of screening policy with the use of short-term testing provides too many false negatives and false positives at and around the guideline of 4 pCi/l. Perhaps some 50 percent of those testing for radon under the EPA protocols for screening have received wrong information. Short-term screening, as it is currently done, works for the much fewer hotter houses at the expense of the vast numbers of marginal houses at and around 4 pCi/l where the bulk of the health risk is located (Harley, 1990). Furthermore, screening measurements fail to provide a significant margin of safety. Ultimately, both screening and short-term testing fail due to the nature of radon variability in houses. Given that EPA protocols and the RMP program legitimize these problems, they also fail.

Prospects for House Characterization

Testing error can best be addressed through a change of testing policy to house characterization, based upon the use of an integrated annual radon level. This approach eliminates significant policy error by refocusing radon policy on the actual distribution of risk. Interpretation error is limited because the results of characterization may be directly applied to the radon risk charts. Methodological errors are minimal because house characterization is not affected by radon variability and, therefore, sampling error is not an issue. Also, there is no potentially incorrect assumption of a maximal value in the basement, as in screening. While even a year-long measurement offers a significant statistical chance that a house with a radon value at or slightly above 4 pCi/l will measure below 4, it offers the most valid and reliable information available about risk at a reasonable cost and is the most logical approach to radon measurement (Nero et al., 1990).

A shift to house characterization faces serious obstacles, however. The revised 1992 EPA *Citizen's Guide* avoids any discussion of the accuracy or validity of measurements and removes the minimal recommendation for house characterization found in the 1986 guide. While the agency purports to believe in the myth of the quick test, using elaborate excuses to justify its policy (see chapter 12), another reason for the agency's intransigence may simply be the difficulty of admitting that they were wrong. Reissuing new protocols on lengthening the sampling time for measurements might make EPA look foolish and raise complex questions of liability. All short-term tests below 4 pCi/l would have to be repeated in order to identify the 20 percent of false negatives. The radon industry's heavy reliance upon the short-term test, encouraged by EPA, would cause serious disruptions. The repudiation of short-term tests (including expensive continuous monitor tests) would invalidate the bulk of real estate transaction measurements.

Of course, continuing with the present testing policy has its own set of problems. First, EPA's current screening policy has created a major credibility

problem. In many states, EPA and others have found, it is difficult enough to get people to screen for radon, let alone to retest later (GAO, 1989; Evdokimoff and Ozonoff, 1992). As homeowners begin to realize how little they learn from a short-term screening test, there is little motivation to retest or to expend a significant sum of money for mitigation based on relatively unreliable information. Neither can it continue to be argued that the present screening approach is cheaper when it so miserably fails to sort houses by risk and may have resulted in remediation of a substantial number of buildings at sale based upon false positive readings. There is precedent for EPA to revise protocols based on new information, as was done previously to eliminate grab sampling as a screening and follow-up measurement method because of unacceptably high error in sampling. In the instance of the quick test, one would be hard put to argue that the situation is much better.¹³

The real goal of policy should be to motivate accurate house characterization, thereby giving people more reliable information on health effects. EPA's radon policy relies on a screening approach, originally designed to motivate testing, that backfired by asking the wrong questions and failing to provide valid and reliable means of measurement. For a responsive radon policy to exist, the myth of the quick test must be finally laid to rest. In the next chapter, we see that there are also problems with the quick fix.

Notes

1. These tests do not require closed-house conditions and can be made either with a year-long alpha track or electret detector or by repeating a charcoal canister test over each of the four seasons and averaging the results (EPA, 1986).

2. The new version of the *Citizen's Guide* (EPA, 1992f) suggests a second immediate short-term test if the screening result is above 10 pCi/l, or if results are needed quickly. Also, the new guidelines suggest averaging screening and follow-up results rather than taking follow-up results on two lived-in levels of the house.

3. Following screening results between 4 and 10 pCi/l, the 1992 EPA guidelines recommend only a long-term test of greater than ninety days, although there is no explicit recommendation for a full year test. The guide additionally suggests averaging screening and follow-up tests even if they are both short term and ignores taking follow-up tests on two lived-in levels. When screening radon levels are higher than 20 pCi/l (10 pCi/l under the 1992 recommendations), shorter-term testing under closed-house conditions is recommended.

4. There were not really any other options available. Grab samples were fast, but they were known to be inaccurate. Like continuous monitoring, they were extremely expensive because skilled labor was required. The alpha track detector was a patented invention, not available for free use by government. The electret had not as yet been invented.

5. Sampling time refers to the length of time of the test. Three common approaches to sampling radon have been used (George, 1988). Grab sampling collects a sample of air over a short time interval, typically minutes, essentially providing an instantaneous snapshot of the radon or daughters in the building at the time the sample was collected. Because radon levels may vary considerably in a given building over time, grab sampling is unlikely to provide a valid picture of the long-term average radon value. Continuous sampling uses sophisticated monitors to repeatedly sample air, typically averaging over one-half to one-hour intervals, and provides detailed information on the variation of radon over time. This method is limited by complexity and cost. Finally, integrated sampling uses detectors to average radon or daughter concentrations over a few days to a year or more, giving a single number, the average, as the result. The commonly used devices, the alpha track detector, the charcoal canister, and the electret, are based on integrated sampling. We focus on integrated measurements in the text.

6. Interestingly, when three-month alpha track results were also compared to yearly averages, a 50 percent uncertainty was found. Thus, the three-month sample is considerably better, but hardly good, at approximating annual averages.

7. One remedy for addressing the particularly alarming number of false negatives—alarming because they falsely reassure the building occupants—would be to drop the radon action level to 2 pCi/l from 4 pCi/l. Under these circumstances, false negatives might drop to 5 percent, but false positive occurrences would become correspondingly more frequent. Steck suggests that a more reasonable corrective would be to eliminate screening measurements altogether.

8. EPA guidances from 1986 to 1992 called for measuring the radon level in the lowest potentially livable area and, after 1992, measuring in the lowest lived-in area.

9. Typically, two-thirds of the measurements made with the same type of detector under the same conditions and time interval would be expected to give a result within plus or minus 10–30 percent of the result obtained. This statistical error in percent is estimated by dividing the square root of the number of counts by the number of counts and multiplying the entire quantity by one hundred.

10. A radon chamber is simply an air tight box in which a known radon source produces a measured radon concentration. The chamber allows calibration factors to be determined for detectors, or it may be used to check the ability of a company to measure the radon level in the chamber, presumably unknown to them.

11. It is interesting that when EPA decided to get a more accurate distribution of actual radon exposure to the population, they took year-long measurements (see chapter 5). Yet when it comes to individual homeowners making expensive decisions about mitigation, the agency considers the short-term test to be adequate.

12. Radon is unlikely to be concentrated in water when the building draws its water supply from a large municipal source, since reservoirs will lose their radon to the outdoor air and large systems based on groundwater take so long to deliver water that most radon is decayed.

13. Given evidence of radon variability and an unacceptable rate of false negatives, it would be prudent to increase sampling times for other measurement methods. Indeed,

in December of 1990, the EPA revised its screening and follow-up protocols for testing to a minimum of two days. Although it is a step in the right direction, this clearly doesn't go far enough. Scott (1988b) and Sextro (1990) suggest something on the order of fifteen to thirty days in order to average over recurrent weather cycles. While still having some of the limitations of a quick test, the two-week to one-month measure may be the shortest duration having any possible claims to validity.