Ecological & Biological Sciences

Exponent®

Critical review of radon testing evaluations by Neri and "The EARTH Study" by Kitto et al.

Exponent®

Critical review of radon testing evaluations by Neri and "The EARTH Study" by Kitto et al.

Prepared for

The National Multifamily Housing Council
The Mortgage Bankers Association

Prepared by

Exponent
15375 SE 30th Place,
Suite 250
Bellevue, Washington 98007

August 5, 2021

2106100.000 - 4699

Contents

Acronyms and Abbreviations **iii**

Executive Summary **iv**

Introduction **1**

Methodological Criticisms of the EARTH Study **3**

The EARTH Study radon data are not nationally representative 3

The EARTH Study does not account for measurement error 4

The EARTH Study cost-benefit and risk analyses are incomplete 6

Conclusion **9**

References **10**

Appendix: Implications of Measurement Error for Decision Making **11**

Acronyms and Abbreviations

ACD	Activated Charcoal Detector
ATD	Alpha-Track Detector
EARTH	Evaluating and Assessing Radon Testing in Housing
EIC	electret ion chambers
L	Liter
NRC	National Research Council
pCi	picoCuries
U.S. EPA	United States Environmental Protection Agency
U.S.	United States

Executive Summary

Exponent conducted a critical review of the study, "Evaluating and Assessing Radon Testing in Housing with multifamily federal financing (the EARTH Study)," by Kitto et al. (2021), as well as a predecessor article, "Evaluation of percentage-based radon testing requirements for federally funded multi-family housing projects," by Neri (2019). For buildings with varying numbers of ground-floor units, the Neri article presents a probability analysis of the sample sizes required to ensure with high probability that testing identifies at least one unit with elevated radon (at or above 4 pCi/L) when such levels are present in the building, given stated assumptions and with acknowledged limitations. Neri offers these findings as "a starting point for a discussion" and concedes that "much work remains to be done to clarify or improve existing radon testing recommendations."

The EARTH Study cites Neri's work and applies the same probability model, but this subsequent evaluation differs in at least three important respects. First, the EARTH Study uses selected data from actual radon testing at multifamily properties. Second, rather than proceeding from the objective that the test sample size be sufficient to detect (with high probability) elevated radon when present, as is the case of the analysis by Neri, the EARTH Study authors impose the stricter requirement that initial testing include (with high probability) all ground-floor units with elevated radon. Third, the EARTH Study authors reach a definitive concluding recommendation regarding radon testing: "[f]or the vast majority of multifamily building sizes, all ground floor units in multifamily buildings should be tested for radon."

Exponent's review finds that the EARTH Study fails to provide much of the clarification called for in the Neri article. Specifically, because of its requirement for exhaustive sampling, the EARTH Study does not consider and provides no additional guidance on such questions as which units to test, whether to test multiple structures on the same property, and how to respond (i.e., with further testing or mitigation) to a measured radon concentration at or above 4 pCi/L.

Furthermore, we have identified multiple methodological issues—including some limitations that are acknowledged but not consistently respected or addressed by the authors of the EARTH Study. These identified methodological issues substantially constrain the reliability of the EARTH Study recommendation for 100% radon testing of ground floor units in multifamily buildings:

- ***The radon data analyzed by the EARTH Study are not representative of nationwide multifamily housing units, which limits the generalizability of study findings***. Non-random data "preferences" were used when compiling data, the compiled data were not geographically or regionally representative, and data did not capture daily or seasonal variation in radon concentrations. The report's estimates of the percentage of units with elevated radon levels missed by testing fewer than 100% of ground-floor units may be inaccurate for areas of the country poorly represented by the data (i.e., U.S. EPA Radon Zone 3), and these inaccuracies would be propagated if the cost-benefit and health risk calculations reported in the EARTH Study were applied at the national scale.

- ***The EARTH Study and Neri analyses fail to consider and properly account for measurement error inherent to radon testing methods, including the risks and associated costs of making incorrect decisions.*** Consequently, the reported analyses are insufficient to support the EARTH Study's finding that 100% ground-floor testing of buildings with up to 20 ground-floor units would provide 95% confidence that no ground-floor units in tested buildings have radon levels exceeding 4 pCi/L. Specifically, 100% sampling will generate significant false positives, and lead to potentially substantial unnecessary mitigation costs that are not included in the EARTH Study cost-benefit analysis. Additionally, the authors do not address the false negatives that occur, even with 100% ground-floor testing, when tests of units with actual radon levels exceeding 4 pCi/L yield measured radon concentrations below 4 pCi/L.
- ***Health cost and risk assumptions are overly simplified, inadequately supported, and not demonstrably applicable to the national population of multifamily housing occupants.*** Estimates of the number of lives saved per decade by mitigating residential units with elevated radon measurements have been applied by the report's authors to a study in which almost half (43%) of the properties were assisted living facilities and did not consider resident demographics, duration of tenancy, hours per day spent indoors, the relative contribution to lifetime radon exposure, or any other confounding factor (e.g., smoking, occupational exposure) affecting lung cancer risk.

Introduction

Exponent has performed a critical review of the study, "Evaluating and Assessing Radon Testing in Housing with multifamily federal financing (the EARTH Study)" by Kitto et al. (2021) and "Evaluation of percentage-based radon testing requirements for federally funded multi-family housing projects", a theoretical study by Neri (2019) on which the EARTH Study is based, as well as supporting and associated documents and reports. We identify limitations of the EARTH Study, in view of the guidance provided by Neri for radon testing in multifamily housing, as well as additional methodological deficiencies in the EARTH Study.

The Neri article presents a probability analysis of the sample sizes required to ensure with high probability that testing identifies at least one unit with elevated radon when such levels are present (at or above 4 pCi/L) in buildings with varying numbers of ground-floor units, given stated assumptions and acknowledged limitations. Of the assumptions underlying this analysis, two are particularly worthy of note:

- *Adaptive sampling*— "[i]dentification of one unit as high radon would result in either further testing of all units or installation of a radon mitigation system for the structure."
- *Statistical independence*— "[r]adon concentrations in each unit of a multifamily housing complex are unrelated."

Neri acknowledges that radon concentrations in adjacent units may be correlated, and the sample sizes estimated in his analysis may therefore be conservative, requiring more testing than necessary to detect the presence of elevated radon at a building. In acknowledgment of the limitations that these assumptions place on the analyses, rather than making specific recommendations on the level of sampling, Neri offers his findings as "a starting point for a discussion" and concedes that "much work remains to be done to clarify or improve existing radon testing recommendations." Neri calls for additional research to inform decisions about which units in a building to test, whether multiple structures on the same property should be tested, and which actions should be taken when a radon concentration above 4 pCi/L is measured during testing of a structure.

Although the EARTH Study cites Neri's work and applies the same probability model, this subsequent evaluation differs in at least three important respects. First, the EARTH Study uses selected data from actual radon testing in multifamily buildings. Second, rather than proceeding from the objective that the test sample size be sufficient to detect (with high probability) elevated radon when present, the EARTH Study authors impose the requirement that initial testing at a building be exhaustive—i.e., that the test sample include (with high probability) all ground-floor units with elevated radon. Third, citing their analyses as the basis, the EARTH Study authors reach a definitive conclusion: "[f]or the vast majority of multifamily building sizes, all ground floor units in multifamily buildings should be tested for radon."

Exponent's review finds that the scope of the EARTH Study fails to provide much of the clarification called for in the Neri study. One key area of uncertainty identified by Neri was the correlation of radon concentrations among units in a building. Although the EARTH Study data from multifamily buildings in which all ground-floor units were tested provide an empirical

basis to estimate this correlation of radon concentrations in adjacent units, the EARTH Study does not explicitly address this aspect or its implications for the radon test sample sizes determined by Neri. Additionally, because of its requirement for exhaustive sampling, the EARTH Study does not consider adaptive or targeted sampling approaches and provides no additional guidance on such questions as which units to test, whether to test multiple structures on the same property, and how to respond (i.e., with further testing or mitigation) to a measured radon concentration at or above 4 pCi/L.

Methodological Criticisms of the EARTH Study

In addition to our concerns about the adequacy of the EARTH Study's scope, we have identified multiple methodological issues—including some limitations acknowledged, but not consistently respected, by the authors—that substantially constrain the reliability of the EARTH Study recommendation for 100% radon testing of ground floor units in multifamily buildings:

- The EARTH Study radon data are not nationally representative;
- The EARTH Study does not account for measurement error;
- The EARTH Study cost-benefit and risk analyses are incomplete.

The technical details and implications of these methodological issues are described and discussed below.

The EARTH Study radon data are not nationally representative

The EARTH Study authors acknowledge several limitations and areas of potential bias in their data set and they take care to note that, "[i]t was not the intent of this study, nor do the authors suggest, that the radon prevalence found in this study is nationally representative of radon at United States (U.S.) multifamily properties." Some acknowledged potential biases include "preferences"[1] for data that included buildings with units with elevated radon data and a preponderance of unit data collected in United States Environmental Protection Agency (U.S. EPA) Radon Zones 1 and 2 (counties with predicted average indoor radon screening levels >4 pCi/L and 2 to 4 pCi/L, respectively), and from 18 states, particularly from Ohio and Illinois. These data preferences explain the greater prevalence of units containing elevated radon levels and higher mean concentrations of radon in the EARTH Study compared to national averages.

The authors of the EARTH Study note that their radon database is, "one of the nation's largest" as justification for proceeding with analyses despite known biases. However, other large radon studies exist that the authors could have used to better understand and address the representativeness of their data. Although the authors reference an earlier national survey of radon levels in homes by the U.S. EPA (Marcinowski et al. 1994), for a minor point regarding data distribution, they do not acknowledge that the instrumentation used in this earlier study had greater accuracy, longer durations of sampling, and more realistic representation of all U.S. EPA zones than those of the EARTH Study. It would be appropriate for the authors to compare their data with results obtained in this national sample of residences to inform the representativeness (and accuracy) of the EARTH Study data, which the authors acknowledge is limited. A comparison of the compiled EARTH Study data with other datasets could highlight the degree to which the recommendations of the EARTH Study should be restricted by the limits of their data, rather than assuming recommendations are appropriate nationwide.

[1] The EARTH Study authors do not describe how a "preference" for certain data influenced their data collection or processing.

The Neri study recognized the importance of radon prevalence in determining the sampling rates required to reach 95% probability that the test sample for a building would include at least one unit with a high radon concentration. Although the Neri study addressed a less demanding testing objective, i.e., ensuring a high probability that the test sample had at least one unit with an elevated radon concentration, the EARTH Study authors should have been aware that sampling designs based on the hypergeometric distribution would require different sampling rates at different levels of radon prevalence to achieve the same testing objective, rather than adherence to a single recommended rate of sampling.

Despite the EARTH Study authors' recognition that their data are not nationally representative of radon levels, they nonetheless erroneously conclude that, "[f]or the vast majority of multifamily building sizes, all ground floor units in multifamily buildings should be tested for radon." The expansive conclusion that 100% sampling is required for all multifamily housing in the United States is not adequately supported for multiple reasons, including the selection biases manifested in the compiled EARTH Study data. The authors' error—not adhering to the limits of the available data—is propagated in subsequent analyses of cost-benefits and risk, and exacerbates the problem of measurement error, as described below.

The EARTH Study does not account for measurement error

The EARTH Study's analysis presumes that no incorrect decisions about the presence of elevated radon will be made if 100% of ground-floor units are tested (see, e.g., Figure 6). In a statistical analysis focused on measuring the probability of omissions (i.e., Table 13) or decision errors (i.e., Table 14) when relying on sample data, it is important to consider the influence of measurement error. In this context, measurement error refers to the extent to which test instruments may incorrectly gauge radon concentrations. Assessing measurements to be "reliable" according to some metric,[2] however, is not sufficient to ensure a particular application of the test method will be robust, i.e., insensitive to measurement errors. Even low error rates from a "reliable" method can yield large numbers of errors when testing is conducted thousands of times at the national scale.

Neither the Neri article nor the EARTH Study substantively addresses measurement error.[3] The Neri study does, however, conclude that radon testing accuracy requires "further analyses" if testing based on hypergeometric results were to be implemented in a "substantial proportion of multi-family housing in the U.S." Subsequently, the EARTH Study made just such a recommendation for widespread testing in a substantial portion (i.e., 100%) of multifamily housing in in the U.S., without any consideration of testing accuracy. The failure to consider measurement error is a serious oversight in the context of the EARTH Study, because of its authors' unqualified advocacy for substantially expanded sampling and testing at the national level. The following discussion and analyses describe sources of measurement error and the

[2] The EARTH Study calculated a "reliability ratio" for duplicate measurements and found it to be "very high" (p. 17).

[3]The Neri article also displays anomalous results in its Figure 1, indicating the probability of identifying units with elevated radon is greater when a smaller percentage of units (10% vs. 25%) is included in the test sample (and high radon occurs at a prevalence of 1 in 15 units). Such a result conflicts with basic principles of statistical inference.

effects of measurement error on rates of false indications and how the number of false indications increases with sample size.

Several devices and protocols are available for measuring the level of radon gas in residential dwellings. All measurement devices have an inherent degree of accuracy that relates to the ability of the device to correctly detect radon concentration as being above or below the 4 pCi/L action level threshold. Some instruments provide greater accuracy than others. The majority (88%) of the EARTH Study radon measurements used activated charcoal detectors (ACDs), while the remaining measurements were made with other devices: alpha-track detectors (ATDs), electret ion chambers (EIC), and liquid scintillation. The typical measurement uncertainty for ACDs is 10-30%, ATDs is 10-25%, and EICs is 8-15% at radon levels of ~5.4 pCi/L.[4] An interlaboratory comparison study utilizing data from over 10 years of testing found that the systematic measurement error of most instruments issued by professional laboratory services can vary ±10% from the true radiation dose values and that a single dose measurement may have an additional random measurement error of ±15% at high dose concentrations At lower dose levels of measured concentrations may vary from the actual level by greater than 50% (Beck et al. 2013).

These estimates of radon measurement error are consistent with those reported in the EARTH Study. As noted in the study report, more than 10% of units with side-by-side duplicate measurements with one value ≥ 4 pCi/L (16 of 133 units) had the other value < 4 pCi/L. From data on 932 units with duplicate radon measurements, the reported within-unit variability was 0.06—or, equivalently, a standard deviation of 0.245 in the log radon scale. This level of measurement error implies, for example, that testing of a unit with an actual radon level of 4 pCi/L could plausibly produce a measured value as low as 2.5 pCi/L or as high as 6.5 pCi/L.[5] Thus, the inherent imprecision of the measurement devices creates the potential for erroneous decisions when they are made by judging the value obtained from an individual test against a numerical threshold.

An additional source of measurement error is the duration of the testing protocols available to the EARTH Study. Most data points reported in the study consisted of samples of 2-3 days duration. Because radon levels fluctuate significantly over time, both on daily and seasonal time scales, and by space use, longer term measurements of 3-12 months are preferred to obtain more representative radon concentration measurements,[6,7] as in the national residential U.S. EPA survey (Marcinowski et al. 1994).

The appendix to this report numerically demonstrates how the reported measurement error in the EARTH Study contributes to misclassification (false positives and false negatives) of elevated radon levels, which is further exacerbated by regional differences in predicted radon levels by U.S. EPA zone. Simulations reported in the appendix show that 100% sampling will generate significant numbers of false positives, and contrary to the EARTH Study conclusion, testing

[4] Table 6 (World Health Organization 2009)

[5] These values correspond to the endpoints of a 95% prediction interval.

[6] https://www.canada.ca/en/health-canada/services/publications/health-risks-safety/guide-radon-measurements-residential-dwellings.html#a1. Accessed July 12, 2021.

[7] Section 2.1.3 (World Health Organization 2009).

100% of all ground-floor units will not necessarily provide 95% confidence that no units exceed the 4 pCi/L action level due to the occurrence of false negative results.

False indications in the EARTH Study are premised on a binary decision as to whether the radon level in a tested unit is above or below a threshold of concern. This same binary approach was also the focus of the Neri study. Such an approach neglects information in the distribution of the numerical values of the measured radon concentrations. Other approaches using the actual measured concentrations do exist (e.g., American Society for Quality 2013) and may result in lower probabilities of decision error at reduced levels of sampling. Because these methods do not reduce individual test results to simple binary outcomes, their application may produce radon testing protocols that offer equal or greater risk mitigation with less sampling.

The EARTH Study cost-benefit and risk analyses are incomplete

The cost-benefit analysis of lung cancer risk arising from missed detections of high radon levels without a 100% sampling protocol is perfunctory and incomplete. These analyses omit consideration of the lack of national representation of the radon data used in the study, the presence of measurement error in the available data as described above, and key factors related to lung cancer risk.

The EARTH Study radon risk analysis aims to determine the reduction in lung cancer risk achieved by mitigating the residential units in their database with radon concentrations greater than 4 pCi/L that would have been forgone in the absence of a 100% sampling protocol. This analysis failed to consider other important factors affecting lung cancer risk, such as resident demographics, occupational exposure, duration of tenancy, the relative contribution to lifetime radon exposure, or smoking—perhaps the most critical confounding factor—in the evaluation of lung cancer risk. The EARTH Study uses estimates of "15,400 to 21,800 radon related lung cancer deaths per year" derived from the National Research Council (NRC 1999) but does not acknowledge this estimate is strongly affected by smoking habits, with only 2,100 or 2,900[8] of 11,000 total lung cancer deaths in non-smokers attributed to radon. The NRC also reported that, "Most of the radon-related deaths among smokers would not have occurred if the victims had not smoked." The NRC report made it clear that smoking should be included in the assessment of radon health risk, but this factor was not considered in the EARTH Study.

The EARTH Study's estimates of the number of lives saved by mitigating residential units in the study with elevated radon measurements are based on data from a set of properties in which almost half (43%) were assisted living facilities. The demographics (age, health status), personal histories (including history of smoking and occupational exposures), and daily activity patterns of the occupants typical of the assisted living facilities will play a large role in their lifetime risk of developing cancer. Thus, the estimates of lives saved for units in the EARTH Study data are

[8] The EARTH Study presents these values as a range of uncertainty for radon-related deaths, but this is incorrect and misrepresents the uncertainty in estimates of radon related deaths. The NRC described these two values as alternative central estimates based on different risk models saying, "15,400 *or* 21,800 per year" (emphasis added), not as the range of radon related deaths, "15,400 *to* 21,800" (emphasis added) used in the EARTH Study. The NRC provides a much larger range for uncertainty, "as low as 3,000 or as high as 33,000," radon-related lung-cancer deaths each year.

likely inaccurate because the radon data are heavily skewed toward assisted living facilities, while the applied risk model parameter values derive from more typical resident demographics and occupancy patterns. Moreover, the EARTH Study authors strongly imply that their calculations can be extrapolated nationwide. As discussed above, the EARTH Study data set is not representative of national radon risk, and such extrapolation is not appropriate.

The lung cancer risk analysis in the EARTH Study does not consider nationwide variation in the risk of radon exposure, such as indicated by the U.S. EPA Radon Zones. The EARTH Study analysis implicitly assumes multifamily housing units in all radon zones will have levels of radon exposure and consequent cancer risk similar to the values calculated from their study data. The EARTH Study risk calculations also implicitly assume constant lifetime exposure to radon at levels for multifamily housing that are estimated from data collected disproportionately from testing of assisted living facilities. The risk calculations conducted by the EARTH Study are based on a series of explicit and implicit model assumptions, some of which are acknowledged, such as spatial and temporal variation, time spent indoors, and particular values of conversion coefficients. Despite recognizing these sources of variation, the EARTH Study authors provide no analysis of the sensitivity of their radon risk calculations and include uncertainty only in the cost of medical treatment for lung cancer. The focus on point estimates in these risk calculations does not allow a full consideration of the variability and evaluation of the relative importance of factors affecting exposure. A probabilistic analysis would be more realistic and useful in a reevaluation of current radon testing practice.

The EARTH Study cost-benefit analysis also assumes that the principal cost associated with 100% sampling is the cost of testing, estimated as approximately $50 per dwelling,[9] and the authors weigh this cost against the number of missed detections without 100% sampling. This cost estimation fails to recognize the cost of the potentially large numbers of false positives and associated costs of unnecessary mitigation measures that can reasonably be expected to greatly exceed $50 per dwelling unit.[10] These unnecessary costs are particularly relevant in areas such as U.S. EPA zone 3, where radon levels are typically lower and Exponent's simulations (see appendix) show up to 25% of positive tests may be erroneous.

The EARTH Study provides an example calculation of the relative cost of a 90% sampling plan compared to a 100% sampling plan for large buildings with 10 or more ground contact units. In the EARTH Study database, these building represent 5,000 total ground contact units, and the EARTH Study authors calculate that a 90% sampling plan would "miss" three units with elevated radon compared to a 100% sampling plan. This estimate of three missed units is likely inaccurate due to measurement errors as discussed above. The authors further estimate that the "cost savings" (i.e., the cost of sampling 10% fewer units) of the 90% sampling plan would be a total of $25,000, or $8,333 for each of the three units missed. The authors compare this $8,333

[9] The EARTH Study uses $50 as a, "fair estimate" of the cost of sampling per dwelling unit. No explanation of how this $50 cost is derived is provided and its accuracy or representativeness is unknown and does not appear to consider the multiple types of testing methods represented in the EARTH Study dataset. This $50 value is repeated herein for comparison purposes only.

[10] The minimum cost of a false positive would be the cost to retest the unit, i.e., the EARTH Study's assumed $50 testing cost.

per unit cost savings to a per unit lung cancer cost of $16,800 to support their argument that 100% sampling is cost effective.

However, in simulated testing of a population having a distribution of radon levels comparable to the units in the EARTH Study, Exponent estimated the false positive rate to be 2.5%. Therefore, for a population of 5,000 units, of which 15% (750 units) have elevated concentrations, sampling 100% of units would be expected to yield 106 (4,250 × 0.025) false positives, while reducing the sampling level from 100% to 90% would produce, on average, 10% fewer positive readings and only 95 (4,250 × 0.9 × 0.025) false positives, reducing the number of misclassified units by 11 (4,250 × 0.1 × 0.025). To demonstrate the sensitivity of the EARTH Study's cost-benefit calculations to measurement error, we calculate that if the cost of unnecessary mitigation associated with the 11 excess false positives expected from increased sampling is greater than $2,310 per unit, then no net benefit would be realized from sampling 100% of ground contact units. Neglecting this consideration, the EARTH Study authors cite their cost analysis as the basis for their conclusion in favor of 100% sampling.

Conclusion

Exponent's review of the recently reported evaluations by Neri and by Kitto, et al. (the EARTH Study) finds that the scope of the EARTH Study fails to provide much of the clarification called for in the Neri article to improve existing radon testing recommendations. Although the EARTH Study data from multifamily buildings in which all ground-floor units were tested provide an empirical basis to estimate the correlation of radon concentrations in adjacent units, the EARTH Study does not explicitly address this aspect or its implications for the radon test sample sizes determined by Neri under an unverified assumption of independence. Additionally, because of its requirement for exhaustive sampling, the EARTH Study does not consider adaptive or targeted sampling approaches to radon testing and provides no additional guidance on such questions as which units to test, whether to test multiple structures on the same property, and how to respond (i.e., with further testing or mitigation) to a measured radon concentration at or above 4 pCi/L.

The EARTH Study's recommendation that 100% of ground-contact units of multifamily housing throughout the United States should be tested for radon is based on an incomplete radon sampling dataset that is not representative of the risk across U.S. EPA Radon Zones and is dependent on a statistical analysis that fails to account for the measurement error inherent in radon testing devices. These shortcomings are compounded in perfunctory analyses of the risks and costs of lung cancer from radon exposure relative to the costs of implementing 100% sampling protocols. Importantly, the EARTH Study—which, to our knowledge, has not been published in a peer-reviewed scientific journal—overlooks the potential for large numbers of false positive indications arising from device measurement error and the greatly expanded testing in areas of low radon risk that are poorly represented in the study data. The cost of these false positives and subsequent unnecessary mitigation could be substantial. A more complete probabilistic analysis of radon risk incorporating these issues, as well as others discussed in this report, would more appropriately characterize the tradeoffs between radon risk, health, and costs.

References

American Society for Quality. 2013. Sampling Procedures and Tables for Inspection by Variables for Percent Nonconforming, ANSI/ASQ Z1.9-2003 (R2013). Quality Press.

Beck, T., Foerster, E., Buchroder, H., Schmidt, V., and J. Doring. 2013. The Measurement Accuracy of Passive Radon Instruments. Radiation Protection Dosimetry, 158(1):59–67.

Kitto, M., Murphy, C., Dixon, S., Wilson, J., Jacobs, D., and J. Malone. 2021. Evaluating and Assessing Radon Testing in Housing with multifamily federal financing (The EARTH Study): FINAL TECHNICAL REPORT and MANUSCRIPT to U.S. HUD for Grant Agreement Number NYHHU0038-17 (p. 30).

Marcinowski, F., Lucas, R., and W. Yeager. 1994. National and regional distributions of airborne radon concentrations in U.S. homes. Health Physics, 66:699–706.

National Research Council. 1999. Health Effects of Exposure to Radon: BEIR VI. National Academies Press.
https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=3375739

Neri, A. 2019. Evaluation of percentage-based radon testing requirements for federally funded multi-family housing projects. Journal of Occupational and Environmental Hygiene, 16(4):302–307. https://doi.org/10.1080/15459624.2019.1566735

World Health Organization (Ed.). 2009. WHO handbook on indoor radon: A public health perspective. World Health Organization.

Appendix: Implications of Measurement Error for Decision Making

The EARTH Study authors undertake a series of statistical analyses to examine the probability that testing ground-floor units at varying levels of sampling will fail to detect the presence of radon levels ≥ 4 pCi/L in one or more units (see, for example, Tables 13 and 14). The implications of measurement error on these analyses when making judgments about individual units on the basis of a single test can be expressed by an operating characteristic curve showing how the probability of a test reading at or above 4 pCi/L will vary depending on the actual radon level in the unit (Figure 1).

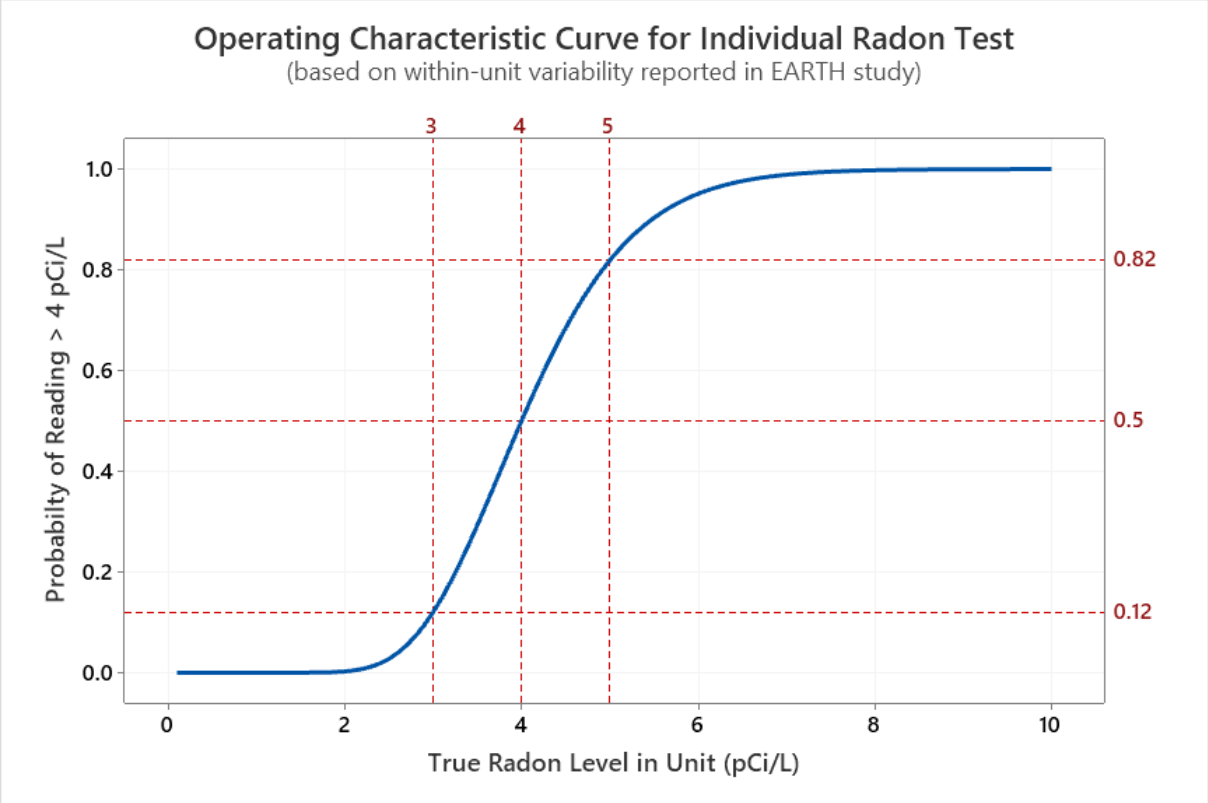

Figure 1. Operating characteristic curve for individual radon tests determined by assuming that the test method is unbiased (in the log radon scale) with an associated standard error equal to the measured within-unit standard deviation (0.245).

In the context of the EARTH Study, test results at or above 4 pCi/L when the true radon level is below 4 pCi/L are false positives; test readings below 4 pCi/L when the true radon level is at or above 4 pCi/L are false negatives. As shown in the above figure, when making a binary decision about whether the unit's radon level is elevated (≥ 4 pCi/L), the reported within-unit variability implies a 12% false positive rate when the true radon level is 3 pCi/L and an 18% false negative rate when the true radon level is 5 pCi/L.

Exponent simulated radon testing of two approximated populations of ground-floor units to demonstrate how applying imprecise test methods on a widespread basis, particularly to such low-risk populations as units in U.S. EPA Zone 3, can generate a considerable number of false indications:

1. A general population with the U.S. EPA-reported average concentration of 1.25 pCi/L and 6% prevalence of units ≥ 4 pCi/L,
2. A low-risk population (corresponding to U.S. EPA Zone 3) with an average concentration of 0.92 pCi/L and 3% prevalence of units ≥ 4 pCi/L (as found in the EARTH Study).

For each of these simulated general and low-risk populations, results from Exponent's simulated tests of 10,000 units are summarized in Table 1 and Table 2 below.

Table 1. Simulated General Population (U.S. average radon concentration and 6% prevalence of units ≥ 4 pCi/L)

Actual	Test < 4 pCi/L	Test ≥ 4 pCi/L	All Tests
Unit < 4 pCi/L	9,247	119	9,366
Unit ≥ 4 pCi/L	71	563	634
All Units	**9,318**	**682**	**10,000**

In the simulation of the results of testing the general population, about one-sixth of all positive test results (119 of 682, 17%) were false indications. The false positive rate was 119/9,366 = 1%, and the false negative rate was 71/634 = 11%.

Table 2. Simulated Low Risk Population (U.S. EPA Radon Zone 3)

Actual	Test < 4 pCi/L	Test ≥ 4 pCi/L	All Tests
Unit < 4 pCi/L	9,606	89	9,695
Unit ≥ 4 pCi/L	46	259	305
All Units	**9,652**	**348**	**10,000**

In the simulation of the low-risk population, about one-fourth of all positive test results (89 of 348, 26%) were false indications. The false positive rate was 89/9,695 = 1%, and the false negative rate was 46/305 = 15%.

The false negative results from these simulations demonstrate that, contrary to the EARTH Study conclusion, testing 100% of all ground-floor units will not necessarily provide 95% confidence that no units exceed the 4 pCi/L action level. For example, in a building with radon concentrations at or above 4 pCi/L in only one unit, that unit has an estimated 18% chance of

being missed, if the unit's true radon level is 5 pCi/L, because of the test yielding a measurement below 4 pCi/L. The probability of decision error at the building level — particularly, missing one or more units with elevated radon levels—will depend on the distribution of radon concentrations among ground-floor units of the same building. As previously noted, the EARTH Study database is not sufficiently representative to provide a reliable national estimate of the probability of a building decision error due to imprecise measurement of radon concentrations.