Add Health
The National Longitudinal Study of Adolescent to Adult Health

Report prepared by

Andrea N. Goodwin

James D. Stewart

Robert A. Hummer

Eric A. Whitsel

Add Health as a Resource for the Science of the Exposome:

Indoor Air

UNC | CAROLINA POPULATION CENTER

CAROLINA POPULATION CENTER | CAROLINA SQUARE - SUITE 210 | 123 WEST FRANKLIN STREET | CHAPEL HILL, NC 27516

https://doi.org/10.17615/ysm0-gh97

This user guide is one in a set of user guides focusing on the built, environmental, and natural features of geopositioned/geocoded Add Health respondent locations over Waves I-VI. Collectively, they describe exposomic measures in the following three domains:

Built Domain	**Environmental Domain**	**Natural Domain**
Commuting Area	Ambient Air	Altitude
Land Use	Indoor Air	Meteorology
Roadway Proximity/Density	Noise	Green space
	Waterborne Lead	Blue space
	Nighttime Light Pollution	
	Solar Irradiation	

Under the Environmental Domain, this particular user guide summarizes the rationale for the estimation of radon concentrations in indoor air. It also documents how the radon concentration source data were acquired, as well as the protocol for quality controlling their assignment and classification across waves.

Acknowledgement

Data for Wave VI of Add Health was supported by two cooperative agreements from the National Institute on Aging (1U01AG071448, principal investigator Robert A. Hummer, and 1U01AG071450, principal investigators Robert A. Hummer and Allison E. Aiello) and a special supplement (U01-AG071450-02S1, principal investigators Robert A. Hummer, Allison E. Aiello, and Eric A. Whitsel) to the University of North Carolina at Chapel Hill. Co-funding for Wave VI was provided by the Eunice Kennedy Shriver National Institute of Child Health and Human Development, the National Institute on Minority Health and Health Disparities, the National Institute on Drug Abuse, the NIH Office of Behavioral and Social Science Research, and the NIH Office of Disease Prevention. Data from Waves I-V of Add Health are from the Add Health Program Project, grant P01 HD31921 (Kathleen Mullan Harris) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, with cooperative funding from 23 other federal agencies and foundations. Add Health was originally designed by J. Richard Udry, Peter S. Bearman, and Kathleen Mullan Harris at the University of North Carolina at Chapel Hill. Add Health is currently directed by Robert A. Hummer; it was previously directed by Kathleen Mullan Harris (2004-2021) and J. Richard Udry (1994-2004). Information on obtaining Add Health data is available on the project website (https://addhealth.cpc.unc.edu).

Citation for User Guide

Goodwin AN, Stewart JD, Hummer RA, Whitsel EA. Add Health as a Resource for the Science of the Exposome**:** *Indoor Air,* Wave VI Data Documentation. Chapel Hill, NC: Carolina Population Center, University of North Carolina at Chapel Hill. https://doi.org/10.17615/ysm0-gh97

.

Table of Contents

1. Introduction

The National Longitudinal Study of Adolescent to Adult Health (Add Health) is a nationally representative sample of U.S. adolescents who were in grades 7-12 during the 1994-1995 school year. Using a complex, school-based cluster-sampling frame, researchers selected high school and feeder school pairs from 80 communities across the United States and drew a sex- and grade-stratified random sample of 20,745 adolescents for inclusion in the study. This sample has been followed from adolescence into early midlife across six waves of data collection to date, with the most recent wave of data collection (Wave VI) taking place between 2022 and 2025 when respondents were ages 39 to 49.

Over the years, Add Health has collected a wealth of information from respondents and their parents about demographic characteristics, familial structures, social relationships, health behaviors, cognition, physical and mental health status, medication usage, and health care access. Add Health also has collected anthropometric, cardiovascular, metabolic, renal, hepatic, inflammatory/immune, infectious, neurodegenerative, and multi-omic biomarkers from respondents. In addition, Add Health has merged multilevel contextual data about the economic, school, neighborhood, policy, and environmental contexts in which the respondents are embedded with the core survey and biological data at each wave. The Add Health dataset thereby provides researchers with rich opportunities to explore the causes and consequences of health status across multiple contextual domains as individuals age across the life course.

This user guide is one in a series documenting the latest contextual and environmental data assembled under the exposome supplement introduced in the preceding acknowledgment. Collectively, the supplemental data and documentation enable researchers to examine a broader array of built, environmental, and natural exposures linked to accurately geopositioned/geocoded Add Health respondent residences from Wave I through Wave VI. Because Wave VI data are not ready for geocoding or dissemination at present, this user guide and the associated data are focused on Wave I-V linkages. The Add Health Team will update this data set and user guide when Wave VI data are available for dissemination.

2. General Overview

The indoor air data include estimated radon concentrations at geocoded Add Health respondent residences. The estimates are partly based on large national surveys including measurements of short-term (2-day), lowest-living-level average radon concentrations and longer-term (1-year), pan-living area average radon concentrations. The rationale for and utility of linking the estimated radon concentrations is described below.

2.1 Rationale

Since its inception, Add Health has continued amassing and disseminating contextual data files across multiple levels of geography, thus resulting in an increasingly comprehensive and diverse set of contextual measures in a nationally representative study spanning adolescence to mid adulthood. In general, these data have been provided to establish infrastructure for research addressing the role of diverse exposures across multiple levels and across the life course in the etiology and disparities of our most pressing health issues. The data collectively position Add Health as a central resource for scientists to more effectively operationalize and study the exposome and its consequences for population health across the life course, with particular attention to disparities across population subgroups.

2.2 Utility

The estimated radon concentrations described herein expand the contextual data available to Add Health researchers, enhancing their capacity to examine the social, environmental, and biological dimensions of the exposome and how they contribute to U.S. population health and disparities. Specifically, they may be valuable to researchers who study health outcomes associated with radon, a ubiquitous, invisible, and odorless gas that is gradually increasing with climate change,[1-4] but also a radioactive and carcinogenic indoor air pollutant that has been associated with risk of cancer,[5-7] hematologic abnormalities,[8-9] as well as cerebrovascular and cardiovascular disease.[10-15]

3. Processing Details

3.1 Estimation and Classification of Radon Concentrations

Indoor radon concentrations originally estimated in 1993 by the US Geological Survey (USGS) and reviewed by the Association of American State Geologists under the Indoor Radon Abatement Act were linked to accurately geopositioned / geocoded Add Health respondent residences from Wave I through Wave VI, as described previously.[10,16] Estimation was based on measurement of short-term (2-day), lowest-living-level average radon concentrations in picocuries per liter (pCi/L);[17-21] architecture type;[22] aerial radioactivity in parts per million of equivalent uranium;[23] geology;[24] and soil permeability.[25] The measurements were scored and summed (yielding a unitless USGS radon index). Then they were categorized (yielding three USGS radon zones, also known as a geologic radon potential)[26-36] and extrapolated to county boundaries (yielding three US Environmental Protection Agency [EPA] radon zones):[37] Zone 3 (<2 pCi/L), Zone 2 (2–4 pCi/L), and Zone 1 (>4 pCi/L), the latter defining EPA's Action Level under its National Radon Action Plan.[38] In the conterminous US, linkage also involved another county-level concentration estimated in a Bayesian hierarchical model of longer-term (1-year), pan-living area average radon concentrations[39] yielding the E.O. Lawrence Berkeley National Laboratory (LBL) county-level geometric mean and related measures.[40-41]

3.2 Assignment of Radon Concentrations

Assignment of the USGS radon index, USGS radon zones / geologic radon potential, EPA radon zones, and LBL measures involved attaching national-level geographic files of the concentrations to geopositioned/geocoded Add Health respondent residences via point-in-polygon intersections. The assignment involved four key activities: (1) identifying, acquiring, and organizing geographic boundary files that most closely matched the temporal and spatial resolution of the source data (e.g., 1990 county boundaries for EPA radon zone data); (2) using ArcGIS/Python software to add source data to the geographic boundary files as polygon attributes and to execute point-in-polygon attribute assignments to respondent locations by Universal Transverse Mercator (UTM) zone; (3) using SAS to aggregate UTM zone-specific attribute assignment files into a national file; and (4) conducting quality control checks to verify the accuracy of the assignments. See Figures 1-3 for maps of the USGS, EPA and LBL source data.

3.3 Quality Control

Given the straightforward nature of point-in-polygon attribute assignments, quality control checks focused on identifying missing or anomalous values by generating basic univariate statistics (n missing;

minimum; maximum; mean) for each radon concentration variable and then confirming them interactively in ArcGIS Pro to ensure that the concentrations matched the corresponding geographic boundary attributes. Concentration assignments were also confirmed for a random selection of points across UTM zones, with special attention paid to geographic areas in which source data were known to be missing (e.g., for LBL measures, no values outside the conterminous U.S.). In addition to confirming concentration assignments for individual respondents, ArcGIS Pro was used to confirm that the overall pattern of individual assignments and missingness matched the overall availability of the source data by state and by larger geographic region.

4. Missing codes

When respondent residential location coordinates were missing and otherwise there was no way to assign possible radon concentrations, -9990 was used to indicate missing radon concentration assignments. The replacement code of -9992 was also used for variables if there were no radon concentration source data available for the time period of interest.

5. Data File

5.1 Structure

The radon data file is provided as a multiple-records-per-respondent long file comprised of 14 variables linked to 49,962 observations. Consistent with other Add Health data, the 20,745 Add Health Wave I respondents are identified by a masked respondent identifier (AID) at every time period during follow-up as presented by the date from (RMEIDADFR) and date to (RMEIDADTO) variables establishing the start and end of each period. Please consult the accompanying codebook for additional details.

5.2 Contents

The radon data file includes the variables below, which are described in the corresponding codebook documentation that also contains frequencies.

Variable Name	**Variable Description**
AID	Add Health Respondent Location ID
RMEIDADFR	Date From
RMEIDADTO	Date To
RMEIDA001	U.S. EPA Radon Zone (1: >4 pCi/L, 2: 2-4, 3: <2)
RMEIDA002	USGS Radon Zone / Geologic Radon Potential (1: >4 pCi/L, 2= 2-4, 3= <2)
RMEIDA003	USGS Radon Index (unitless values of 5-16)
RMEIDA004	LBL Estimated Geometric Mean (GM)
RMEIDA005	LBL Multiplicative uncertainty in GM
RMEIDA006	LBL Estimated Arithmetic mean (AM)
RMEIDA007	LBL Additive uncertainty in AM
RMEIDA008	LBL Est. fraction homes w/ long-term exp. > 4 pCi/L
RMEIDA009	LBL Est. fraction homes w/ long-term exp. > 10 pCi/L
RMEIDA010	LBL Probability > 10% of homes exceed 4 pCi/L

RMEIDA011 LBL Probability > 5% of homes exceed 10 pCi/L

6. References

1. Stanley FKT, Irvine JL, Jacques WR, et al. Radon exposure is rising steadily within the modern North American residential environment, and is increasingly uniform across seasons. Sci Rep 2019;9(1):18472.
2. Ilacqua V, Dawson J, Breen M, Singer S, Berg A. Effects of climate change on residential infiltration and air pollution exposure. J Expo Sci Environ Epidemiol 2017;27(1):16-23.
3. Symonds P, Rees D, Daraktchieva Z, et al. Home energy efficiency and radon: an observational study. Indoor Air 2019;29(5):854-864.
4. Pigg S, Cautley D, Francisco PW. Impacts of weatherization on indoor air quality: a field study of 514 homes. Indoor Air 2018;28(2):307-317.
5. National Research Council. Health Risks From Exposure to Low Levels of Ionizing Radiation: BEIR VII Phase 2. National Academies Press; 2006. doi:10.17226/1134.
6. Ngoc LTN, Park D, Lee YC. Human health impacts of residential radon exposure: updated systematic review and meta-analysis of case-control studies. Int J Environ Res Public Health 2022;20(1):97.
7. Bozigar M, Konstantinoudis G, Zilli Vieira CL, Li L, Alwadi Y, Jones RR, Koutrakis P. Domestic radon exposure and childhood cancer risk by site and sex in 727 counties in the United States, 2001-2018. Sci Total Environ 2024;954:176288.
8. Anthony KM, Collins JM, Love SM, Stewart JD, Buchheit SF, Gondalia R, Schwartz GG, Huang DY, Meliker JR, Zhang Z, Barac A, Desai P, Hayden KM, Honigberg MC, Jaiswal S, Natarajan P, Bick AG, Kooperberg C, Manson JE, Reiner AP, Whitsel EA. Radon exposure, clonal hematopoiesis, and stroke susceptibility in the Women's Health Initiative. Neurology. 2024;102(2):e208055.
9. Shi C, Collins JM, Stewart JD, Williamson MR, Conneely KN, Jung SY, Manson J, Schwartz GG, Reiner AP, Smith RL, Whitsel EA. Indoor radon concentrations are associated with leukocyte and platelet counts. EPI/Lifestyle 2025 Scientific Sessions;P2031. https://aha.abstractarchives.com/abstract/epi2025-4204795/indoor-radon-concentrations-are-associated-with-leukocyte-and-platelet-counts
10. Buchheit SF, Collins JM, Anthony KM, Love SM, Stewart JD, Gondalia R, Huang DY, Manson JE, Reiner AP, Schwartz GG, Vitolins MZ, Schumann RR, Smith RL, Whitsel EA. Radon exposure and incident stroke risk in the Women's Health Initiative. Neurology 2024;102(4):e209143.
11. Buchheit SF, Collins JM, Stewart JD, Williamson MR, Bozeman D, Allison MA, Manson J, Meliker JR, Vitolins MZ, Schwartz GG, Smith RL, Whitsel EA. Indoor radon concentrations are associated with incident myocardial infarction in the Women's Health Initiative. AHA EPI/Lifestyle 2025 Scientific Sessions;MP19. https://aha.abstractarchives.com/abstract/epi2025-4200430/indoor-radon-concentrations-are-associated-with-incident-myocardial-infarction-in-the-womens-health-initiative.
12. Zhang Y, Yan Q, Angley M, Lu L, Miller EC, Judd SE, Field RW, Kahe K. Smoking modifies the association between radon exposure and incident ischemic stroke: The REGARDS Study. Stroke. 2023;54(11):2737-2744.
13. Lu L, Zhang Y, Chen C, Field RW, Kahe K. Radon exposure and risk of cerebrovascular disease: a systematic review and meta-analysis in occupational and general population studies. Environ Sci Pollut Res Int. 2022;29(30):45031-45043.
14. Zhang Y, Lu L, Chen C, Field RW, D'Alton M, Kahe K. Does protracted radon exposure play a role in the development of dementia? Environ Res 2022;210:112980.
15. Kim SH, Park JM, Kim H. The prevalence of stroke according to indoor radon concentration in South Koreans: Nationwide cross section study. Medicine (Baltimore). 2020;99(4):e18859.

16. Buchheit SF, Collins JM, Anthony KM, Love SM, Stewart JD, Gondalia R, Huang DY, Manson JE, Reiner AP, Schwartz GG, Vitolins MZ, Schumann RR, Smith RL, Whitsel EA. Supplemental Appendix. Radon exposure and incident stroke risk in the Women's Health Initiative. Figure S2. Provenance of the radon data. Accessed March 4, 2025. links.lww.com/WNL/D383.
17. United States Environmental Protection Agency. Air and Radiation (6604J). Natl Radon Database Vol 1 State/EPA Residential Radon Survey EPA-402-R-93-001A. 1993. p. PB93-134781.
18. United States Environmental Protection Agency. Air and Radiation (6604J). Natl Radon Database Vol 2 State/EPA Residential Radon Survey EPA-402-R-93-001B. 1993. p. PB93-134799.
19. United States Environmental Protection Agency. Air and Radiation (6604J). Natl Radon Database Vol 3 State/EPA Residential Radon Survey EPA-402-R-93-001C. 1993. p. PB93-134807.
20. United States Environmental Protection Agency. Air and Radiation (6604J). Natl Radon Database Vol 4 State/EPA Residential Radon Survey EPA-402-R-93-001D. 1993. p. PB93-134815.
21. United States Environmental Protection Agency. Air and Radiation (6604J). Natl Radon Database Vol 5 State/EPA Residential Radon Survey EPA-402-R-93-001E. 1993. p. PB93-134823.
22. United States Census Bureau. American Housing Survey. Accessed February 28, 2021. https://www.census.gov/programs-surveys/ahs/data.html.
23. Duval JS, Carson JM, Holman PB, Darnley AG. Terrestrial Radioactivity and Gamma-Ray Exposure in the United States and Canada, 2005. Open-File Report 2005-1413. Accessed June 2, 2022. pubs.usgs.gov/of/2005/1413/index.htm.
24. Geologic Maps of US States. United States Geologic Survey. Accessed November 13, 2022. mrdata.usgs.gov/geology/state/.
25. Principal Kinds of Soils: Orders, Suborders, and Great Groups: U.S. Geological Survey; 1987.
26. Schumann R. Geologic radon potential of EPA Region 1; Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, and Vermont. 1993;Open-File(Report 93-292-D).
27. Schumann R. Geologic radon potential of EPA Region 2; New Jersey and New York. 1993;Open-Fil(Report 93-292-B).
28. Schumann R. Geologic radon potential of EPA Region 3; Delaware, Maryland, Pennsylvania, Virginia, and West Virginia. 1993;Open-File(Report 93-292-C).
29. Schumann R. Geologic radon potential of EPA Region 4; Alabama, Florida, Georgia, Kentucky, Mississippi, North Carolina, South Carolina, and Tennessee. 1993;Open-File(Report 93-292-D).
30. Schumann R. Geologic radon potential of EPA Region 5; Illinois, Indiana, Michigan, Minnesota, Ohio, and Wisconsin. 1993;Open-File(Report 93-292-E).
31. Schumann R. Geologic radon potential of EPA Region 6; Arkansas, Louisiana, New Mexico, Oklahoma, and Texas. 1993;Open-File(Report 93-292-F).
32. Schumann R. Geologic radon potential of EPA Region 7; Iowa, Kansas, Missouri, and Nebraska 1993. 1993;Open-File(Report 93-292-G).
33. Schumann R. Geologic radon potential of EPA Region 8; Colorado, Montana, North Dakota, South Dakota, Utah, and Wyoming. 1993;Open-File(Report 93-292-H).
34. Schumann R. Geologic radon potential of EPA Region 9; Arizona, California, Hawaii, and Nevada. 1993;Open-File(Report 93-292-I).
35. Schumann R. Geologic radon potential of EPA Region 10; Alaska, Idaho, Oregon, and Washington. 1993;Open-File(Report 93-292-J).
36. Schumann R. Geologic radon potential of Guam and Puerto Rico. 1995;Open-File(Report 93-292-K).
37. EPA Map of Radon Zones: National Summary. U.S. Environmental Protection Agency; 1993. Accessed November 13, 2022. epa.gov/radon/epa-map-radon-zones.
38. The National Radon Action Plan 2021-2025. United States Environmental Protection Agency; 2022. Accessed January 25, 2023. radonleaders.org/resources/nationalradonactionplan.

39. United States Environmental Protection Agency. Air and Radiation (6604J). Natl Radon Database Vol 6 National Residential Radon Survey EPA-402-R-93-013. 1993. p. PB93-134831.
40. Lin C, Gelman A, Price P, Krantz D. Analysis of local decisions using hierarchical modeling, applied to home radon measurement and remediation. Stat Sci. 1999;14(3):305-337.
41. Price PN, Nero A, Revzan K, Apte M, Gelman A, Boscardin W. Predicted Parameters of Radon Distributions within Counties in the 48 Conterminous States: E.O. Lawrence Berkeley National Laboratory. Environmental Energy Technologies Division. Indoor Environment Department. The High-Radon Project: Files and Documentation. Accessed June 1, 2022. eetd.lbl.gov/iep/high-radon/ctypred.htm.

Figure 1

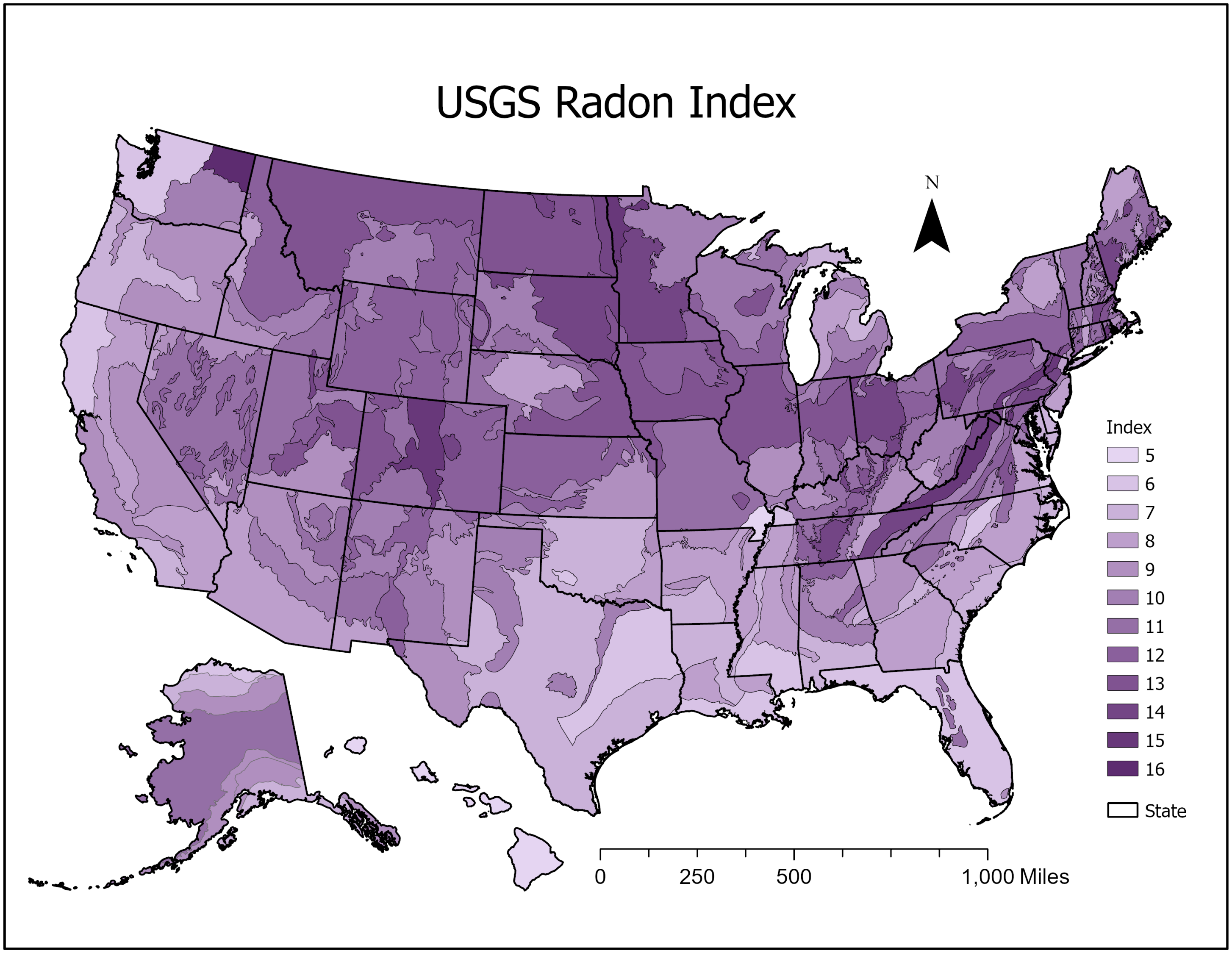

USGS Radon Index
N
Index
5
6
7
8
9
10
11
12
13
14
15
16
State
0
250
500
1,000 Miles

Figure 2

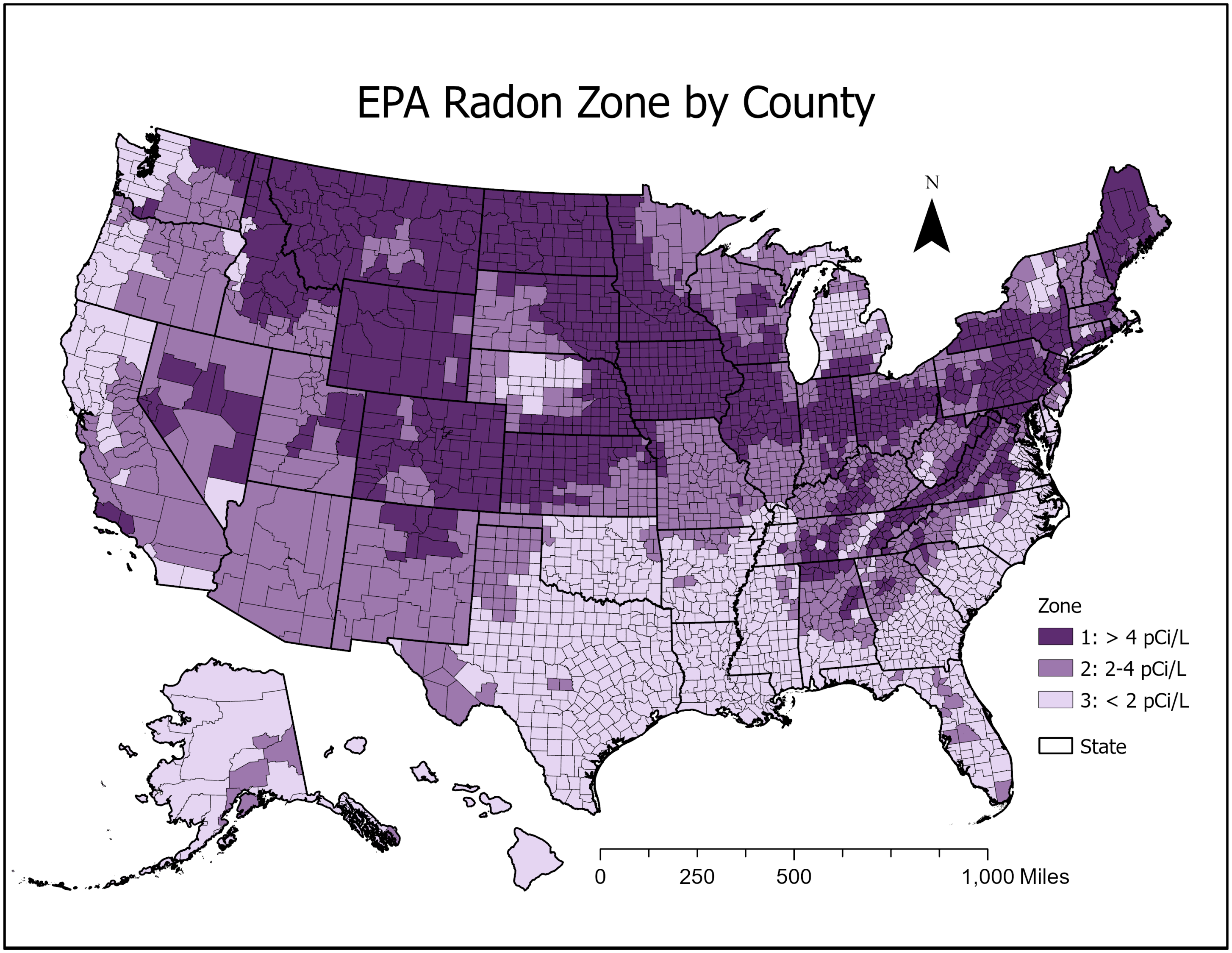
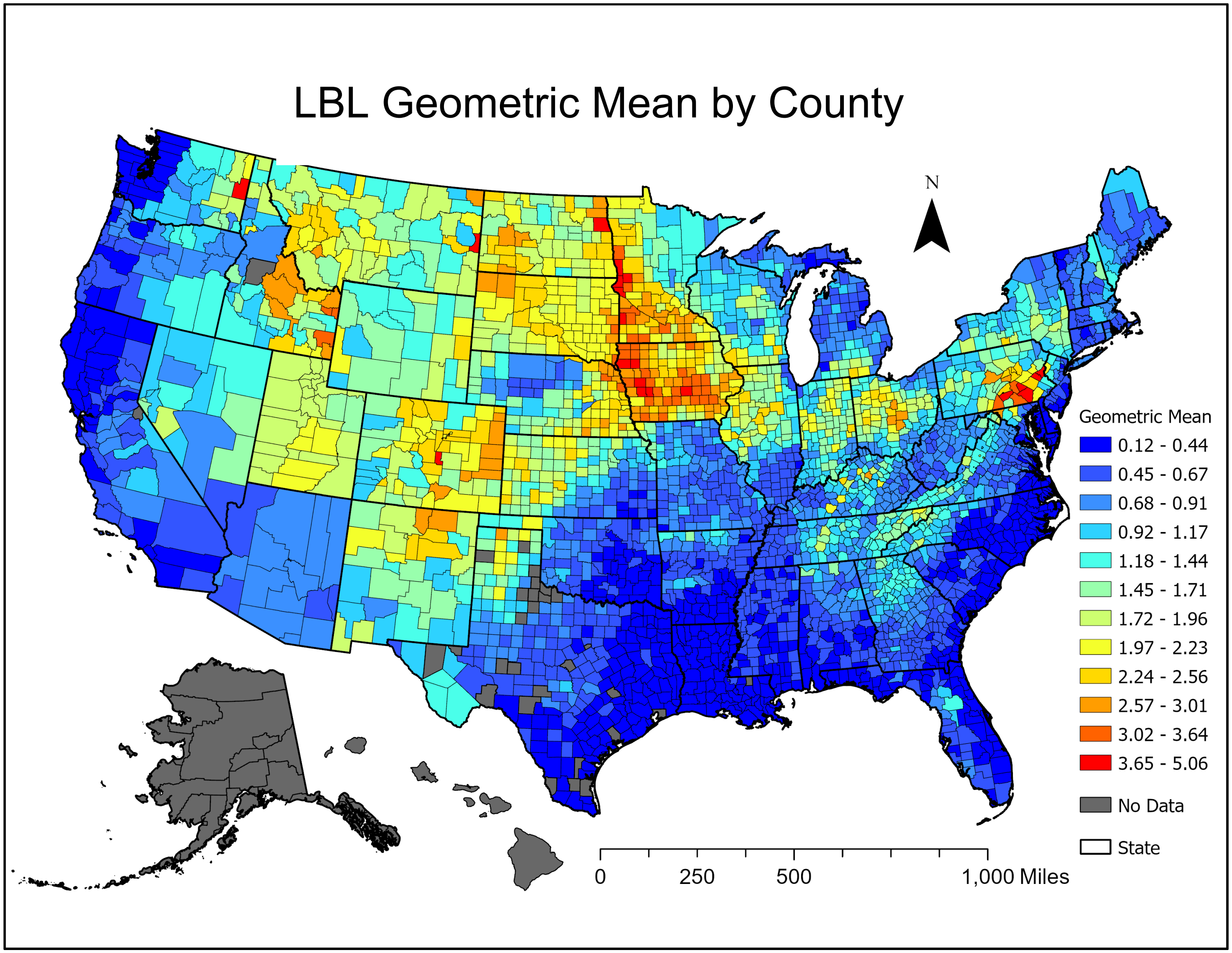

EPA Radon Zone by County
N
Zone
1: > 4 pCi/L
2: 2-4 pCi/L
3: < 2 pCi/L
State
0
250
500
1,000 Miles

Figure 3



LBL Geometric Mean by County
N
Geometric Mean
0.12 - 0.44
0.45 - 0.67
0.68 - 0.91
0.92 - 1.17
1.18 - 1.44
1.45 - 1.71
1.72 - 1.96
1.97 - 2.23
2.24 - 2.56
2.57 - 3.01
3.02 - 3.64
3.65 - 5.06
No Data
State
0
250
500
1,000 Miles