# EVALUATING THE HEALTH EFFECTS OF LONG-TERM RADON EXPOSURE AMONG THE VETERAN POPULATION

By

Lachlan Watkins

## 0.1 Introduction

This paper aims to contribute to the existing literature on the health effects of indoor radon exposure by examining the relationship between cumulative radon exposure and the probability of developing lung cancer among United States veterans. While extensive research has been conducted to explore the association between radon exposure and lung cancer in the general population, limited evidence exists regarding the effects of radon among specific subpopulations, including veterans. Furthermore, lung cancer is the leading cause of death among veterans. Consequently, this study aims to bridge this gap by investigating the heterogeneous effects of cumulative radon exposure within this underserved population. By accurately quantifying the causal effect of residential levels of cumulative indoor radon, this study offers valuable insights into the potential widespread health implications of radon exposure for the overall veteran population.

While there is a considerable body of literature investigating the impact of radon exposure on lung cancer within residential settings, this research has faced various challenges. These challenges include limitations in data availability, measurement errors, utilization of non-random case-control study designs, and concerns regarding external validity. In an ideal research scenario, indoor radon values would be randomly assigned to the study population and continuously monitored throughout each participant's lifetime, allowing for the utilization of cross-sectional and time-series variation to identify causal effects. However, such an approach raises ethical concerns. An alternative approach involves leveraging exogenous shocks to long-term indoor radon levels, but this requires measurements of long-term indoor radon to identify underlying structural relationships. Overcoming these challenges is further complicated by the hypothesis that residential levels of indoor radon exposure have relatively small marginal effects on the probability of developing lung cancer, particularly when compared to factors like smoking or exposure to high levels of occupational hazards. As a result, precise estimates of these effects necessitate a sufficiently large sample size.

This paper aims to tackle these challenges by leveraging the predictive models developed in Watkins (2024). It combines administrative and claims data from the Veteran Health Administration (VHA) with residence information obtained from the ZTRAX database and long-term weather data. This unique dataset, along with the predictions of atmospheric bismuth-214 and indoor radon from the models of Watkins (2024), provides a practical approach for estimating both reduced-form and structural models that examine the relationship between radon risk and the probability of developing lung cancer.

Let's consider the reduced-form relationship between predicted atmospheric bismuth-214 (bi-214)[1] and the risk of lung cancer. Regardless of the origin of radon-222, which produces bi-214, a significant association between atmospheric bi-214 and lung cancer development would suggest that radon-222 progeny contributes to lung cancer risk. However, it is important to acknowledge the limitations inherent in any reduced-form

[1] Bi-214 is a radioactive daughter isotope of radon-222 that releases dangerous ionizing alpha radiation when it decays.

analysis. Outdoor levels of bi-214 and radon-222 generally represent low concentrations of exposure to environmental radiation. Since the primary driver of lung cancer risk from radiation comes from elevated **indoor** exposures, it is unlikely that atmospheric (outdoor) bi-214 plays a major role in lung cancer development. This is particularly relevant considering that most individuals spend the majority of their time indoors. Therefore, atmospheric bi-214 is likely capturing the effects of elevated indoor concentrations. Furthermore, interpreting the marginal effects of outdoor bi-214 on lung cancer from a health policy perspective presents challenges: what actionable measures can be taken to reduce the impact of natural background radiation in the atmosphere on the development of lung cancer?

By providing both reduced-form and structural estimates, this study effectively addresses the potential limitations associated with relying solely on a reduced-form analysis. The individual fixed-effects models for both cumulative predicted atmospheric bi-214 and cumulative predicted indoor radon yield statistically and economically significant effects on the probability of developing lung cancer within this specific population. Notably, using natural log transformations of atmospheric bi-214 and indoor radon produces almost identical marginal effects.

At the mean lung cancer risk, a 10 percent increase in atmospheric bi-214 or indoor radon leads to a 6.2 percent increase in the probability of developing lung cancer. Moreover, increasing cumulative radon by 50 percent results in a substantial 26 percent increase in the probability of lung cancer at the mean level. These effects are remarkably large, especially when considering the predicted levels of long-term radon exposure. In fact, the mean yearly indoor radon concentration is only 1.63 piC/L, and the 99th percentile reaches 3.89 piC/L. Even at these relatively low levels, indoor radon exposure demonstrates a significant impact on the probability of developing lung cancer.

The results of this study yield fascinating insights when exploring the heterogeneous effects of radon exposure among different subpopulations of veterans. Specifically, separate models are estimated based on smoking status, combat history, and veterans diagnosed with chronic obstructive pulmonary disease (COPD). The analysis focuses on estimating the probability of lung cancer and death as a function of cumulative radon exposure and approximating the fraction of lung cancer deaths attributed to radon exposure.

The findings reveal that veterans diagnosed with COPD face the highest risk of dying from radon-attributed lung cancer, with 35 percent of lung cancer deaths among COPD-diagnosed veterans attributed to radon exposure. Similarly, smokers exhibit high levels, with the radon-attributed fraction (RAF) hovering around 28 percent. Notably, there is a distinct pattern wherein populations at greater risk or those considered ”sicker” demonstrate larger RAFs compared to their low-risk or ”healthier” counterparts. This finding may appear surprising at first, as one might expect the radon effects relative to death risk to be higher among the healthier population. However, it is important to consider the levels of radon exposure across the entire cohort

of veterans. Approximately 99 percent of these veterans have yearly predicted levels below 3.89 piC/L, and 75 percent have levels below 1.84 piC/L, well below the suggested remediation threshold of 4 piC/L. Given these relatively low levels of yearly radon exposure, it is unlikely that cumulative exposure from radon alone would generate lung cancer deaths among populations that are not already at an increased risk of lung cancer.

Understanding the heterogeneous effects of radon exposure among veterans holds significant public health implications. By identifying subpopulations with higher susceptibility to radon-induced lung cancer, targeted interventions and preventive strategies can be developed. Tailoring efforts towards at-risk groups, such as smokers or individuals with COPD, may prove more effective in reducing the burden of lung cancer. Furthermore, the establishment of different remediation thresholds for these at-risk populations may be necessary.

In conclusion, this paper presents a novel analysis that leverages a unique dataset to estimate the effects of radon exposure on the probability of lung cancer among veterans. The insights gained from this study enhance our understanding of radon exposure risks within the veteran population and provide the first estimates of the causal effects of long-term cumulative radon exposure on the risk of lung cancer at residential levels of exposure among US veterans. The significant risk of lung cancer observed even at relatively low levels of radon exposure, compared to the EPA's suggested remediation threshold of 4 piC/L, suggests the need to develop a new threshold specifically for veterans, particularly those with COPD or smoking history. Future research should involve collaboration between the VHA and EPA to reduce veterans' exposure to indoor radon and alleviate the burden of lung cancer within this population.

## 0.2 Background

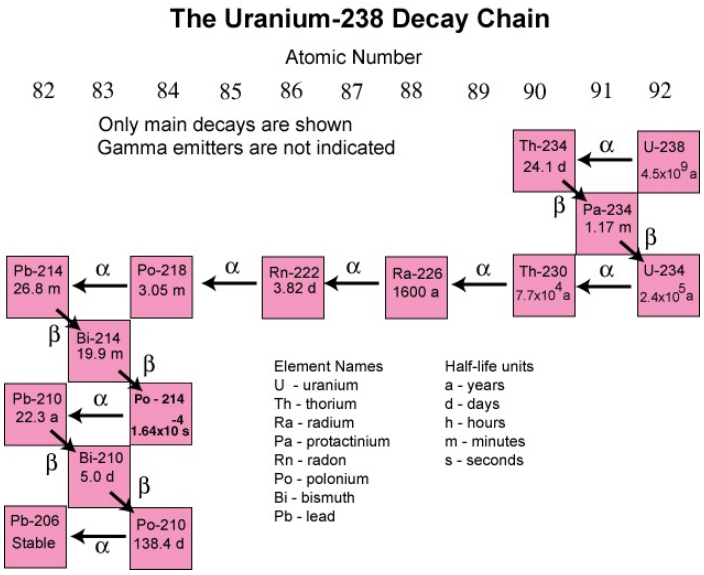
### 0.2.1 Radon and Radioactivity

Indoor radon is a naturally occurring radioactive noble gas that is produced through the decay of radium isotopes. The longest-lived radium isotope, radium-226, has a half-life of approximately 1600 years, indicating that radon and radium originate from the decay of uranium and thorium that were deposited during the formation of the Earth. Figure 1 depicts the decay chain for uranium-238.[2] Radon-222 has a half-life of about 3.8 days, allowing it to diffuse from the ground into homes before emitting harmful alpha radiation and generating statically charged daughter particles that also emit alpha radiation.

It is worth noting the presence of the bismuth-214 isotope in Figure 1, which is part of the direct decay chain of radon-222. Bismuth-214 has a short half-life of about 20 minutes, and the two preceding isotopes (pb-214 and po-218) also have half-lives of less than 30 minutes. Due to the short half-lives of these daughter

[2]Note that thorium-232 also produces an isotope of radon, radon-220, with a half-life of 55 seconds. This short half-life implies that it is unlikely to provide a significant source of indoor radon exposure. Nonetheless, as noted in (Nazaroff, 1992), studies have found that advection is the primary process by which radon enters the atmosphere through uncovered soil. Since the advection of radon-222 will be impacted by the amount of radon-220 that contributes to the advective velocities in the ground, thorium-232 should be included in any model of indoor radon (in other words, thorium-232 shoots out alpha particles, and generates radon-220 down the decay chain, which in turn increases the advective velocities in the soil that may co-determine the exhalation rates of radon-222).

Figure 1: The Radioactive Decay Chain of Uranium-238




isotopes of radon-222, bismuth-214 is an effective measurement for estimating the concentration of radon gas in the immediate vicinity.[3] This observation is critical and serves as the primary predictor of indoor radon levels.

Suppose we observe a deposit of radium-226 in a closed setting. When the observation period begins, the concentration of radon-222 will instantaneously increase as individual radium-226 nuclei randomly disintegrate into radon-222.[4] In a closed setting, radon gas will reach a steady state within 5-7 days, during which new radon-222 is produced from radium-226 at the same rate that current radon-222 decays into polonium-218 (Lehnert and Kearfott, 2010). However, in the underlying rocks and soil, radon-222 has already reached a steady state due to the 4.5 billion-year half-life of uranium-238. Therefore, as radon gas escapes from rocks and soil, it continuously and almost instantaneously produces bismuth-214 along the decay chain, which can be mathematically represented by the Bateman Equation.[5] To see why this is the case, consider the generalized form of the Bateman Equation:

$$X_n(t) = X_1(0) \prod_{k=1}^{n-1} \lambda_k \sum_{i=1}^{n} \prod_{j=1}^{n} \frac{e^{-\lambda_i t}}{\lambda_j - \lambda_i} \tag{1}$$

where $X_1(0)$ is the initial concentration of radon-222, $t$ indexes time, and $\lambda$ index the decay constants (for

[3] In fact, bismuth-214 is one of the isotopes used in gamma-ray counting to determine radon concentrations from indoor activated charcoal radon tests.

[4] One gram of radium-226 decays at a rate of approximately 2.2 trillion decays per second. Radioactive measurements are commonly expressed in picocuries per liter (piC/L) or becquerels per cubic meter (Bq/m$^3$). One piC/L corresponds to 0.037 radioactive disintegrations per second in every liter of air. As a reference, the average human lung capacity is about 6 liters, and one cubic foot of air holds approximately 28.3 liters.

[5] The Bateman Equation describes the activities in a decay chain as a function of time and initial concentrations.

rn-222, po-218, pb-214, and bi-214). Given a positive amount of initial radon-222 in time period 0, the amount of bi-214 at time period 1 (say, 10 minutes later) is strictly positive ($X_n(1)$). For example, suppose radon-222 is constant at ground level at 0.1 becquerels/m$^3$. Then the Bateman equations estimate about 0.002 becquerels of bi-214 would have been produced after 10 minutes in every cubic meter of air. It is important to note that radon-222 is a gas and subject to flow. Indeed, it is possible for radon-222 nuclei that have yet to decay to travel a certain distance before bismuth-214 is produced along the decay chain.[6] This introduces a potential source of measurement error when using atmospheric bismuth-214 as a proxy for local atmospheric radon-222 since the bismuth-214 measurements represent a combination of locally generated and transported radon-222. However, despite this inherent limitation, atmospheric bismuth-214 remains an ideal proxy for atmospheric radon-222 and offers valuable insights into the data-generating process that brings radon-222 to the surface.

### 0.2.2 The Model of Radon

Indoor radon concentrations among the Veteran sample population are unobserved. Thus, to estimate the health effects of cumulative radon exposure on the likelihood of developing lung cancer, I rely on the predicted monthly radon concentrations from the national radon model developed in Watkins (2024). The short-term radon model developed in Watkins (2024) involved generating a complete map of rock lithology for the contiguous United States. This map, combined with geochemical surveys, long-term weather data, soil characteristics, and terrain features, modeled the distribution of radon-accessory elements. Subsequently, it predicted daily atmospheric bi-214, a critical proxy variable for outdoor/surface-level radon. Finally, the predictive model for indoor radon accounts for endogenous building characteristics, sample selection, and uses daily predicted atmospheric bi-214, weather data, and building characteristics to model the data-generating process of short-term indoor radon. Refer to Watkins (2024) for a full discussion of the models.

### 0.2.3 Literature

Despite the risks of radon gas, the economic and health literature on residential radon exposure is limited and relies on small case-control studies and meta-analyses[7]. Darby et al. (1998) performed a case-control analysis in southwest England with 4167 subjects and estimated that the relative risk of lung cancer increased by 0.08 per 100 Bq/$m^3$ (2.7 pCi/L). Field et al. (2000) performed a similar case-control study of female Iowa residents who occupied their current homes for at least 20 years. The study found noisy estimates, predicting excess

[6] In fact, radon-222 is the heaviest naturally occurring gas. As such, radon-222 at ground level will usually form low, flat clouds which spread based on their own density and surrounding terrain. At some point, the heavy gas cloud of radon will dilute, and reach a point where it is no longer heavy as its density becomes close to the surrounding air (Blackmore et al., 1982).

[7] See Lubin et al. (1994),Pavia et al. (2003),Darby et al. (2005),Krewski et al. (2005),Darby et al. (2006),Krewski et al. (2006),Malinovsky et al. (2018),Li et al. (2020), and Chen et al. (2021).

odds ratios of 0.5 for radon exposure at 148 Bq/$m^3$. Darby et al. (2005) analyzed individual data from 13 European case-control studies, finding a significant linear dose-response relationship between indoor radon and lung cancer. They estimate increased relative risks of lung cancer by 0.16 per 100 Bq/$m^3$, twice the effect of Darby et al. (1998). A more recent meta-analysis performed by Malinovsky et al. (2018) analyzed the results of 31 case-control studies from 13 countries and found statistically significant lung cancer risk for radon concentrations above 92 Bq/$m^3$.

The Committee on Biological Effects of Ionizing Radiation (BEIR) VI proposed a model relating radon risk to lung cancer in 1999 using data from 68,000 metal miners from 11 international observational studies. The BEIR VI found significant lung cancer risks for metal miners but relied on limited regression controls to extrapolate the miner results to the general population and did not control for other potential confounding factors. Cao et al. (2017) re-evaluated the estimates from the BEIR VI model while accounting for diesel engine exhaust exposure, reducing the BEIR VI radon-attributed lung cancer estimates by 9-26 percent. This study calls into question the EPA's use of the BEIR VI model to develop the remediation threshold of 148 Bq/$m^3$ (equivalent to 4 picocuries per liter and the EPA's Zone 1 radon designation). Other studies have also highlighted the uncertainty regarding the effects of residential radon exposure and lung cancer risk. Wilcox et al. (2008) performed a case-control study in New Jersey involving 561 cases and 740 control subjects, finding no statistically significant effect of indoor radon on lung cancer risk. Sandler et al. (2006) analyzed 1474 lung cancer cases against 1811 telephone-screened controls from Connecticut, Utah, and southern Idaho. The study's estimates of the excess relative risks for lung cancer weren't statistically significantly different from zero. Another case-control study found no significant risk for lung cancer among non-smoking women in Germany using 234 female lung cancer patients and 535 population controls (Kreuzer et al., 2002). It is clear that the literature on residential lung cancer contains contradictory and uncertain evidence regarding the health effects of indoor radon. Non-random case-control designs combined with small sample sizes lack the power to detect the effects of indoor radon on health outcomes for the residential population across the US.

### 0.3 Data

Demographic, socioeconomic, health history, and address data for veterans were obtained from the Corporate Data Warehouse (CDW) hosted by the VA Informatics and Computing Infrastructure (VINCI). The Institutional Review Board approved the study to include all veterans followed by the Veteran Health Administration (VHA) from 2002 to the present. However, address histories were not available in the CDW database, requiring additional data requests within VINCI. Considering data and computational constraints, the analysis focused on veterans residing in Alabama, Arkansas, Kentucky, Louisiana, Mississippi, Oklahoma, and Tennessee. The veteran address files contained information on the date of address updates, with

only the most recent address included. While some address changes may represent confirmations rather than actual changes, the assumption was made that the update date indicates when a veteran started living at the address on file, which determined their inclusion in the analysis. These criteria resulted in a sample of 1.317 million veterans across the seven states.

Data on veterans' demographics, socioeconomics, health histories, and addresses come from the Corporate Data Warehouse (CDW) hosted by the VA Informatics and Computing Infrastructure (VINCI). The Institutional Review Board application approved the study to include any veterans that were followed by the Veteran Health Administration (VHA) in 2002 to the present. The veterans' address histories are currently unavailable within the CDW database and require additional data requests inside VINCI. Additionally, data and computational restrictions would make analyzing the full US veteran population beyond the scope of this paper. Given these considerations, I limited the analysis to veterans that reside in Alabama, Arkansas, Kentucky, Louisiana, Mississippi, Oklahoma, and Tennessee.

The veteran address files contain information on the date for which a home address was updated and only contain the most recent address. An address could be updated if an employee confirmed an address or altered an address in any way. It is unclear if some address change dates simply reflect an employee confirming a patient's address, rather than an actual change of address. Regardless, I assume that the address update date indicates when a veteran has started living in the address on file, and use this date to determine when a veteran enters the analysis. These restrictions result in 1.317 million veterans across the seven states.

To match veterans to their residences, the addresses were compared to the Zillow ZTRAX database. The matching algorithm prioritized minimizing false matches, resulting in a match rate of 71 percent and 936K matched veterans.[8]. Restricting the sample further to veterans living in single-family homes yielded a cohort of slightly over 700K veterans.

Demographic and health indicators were constructed using the administrative and claims data available in the Corporate Data Warehouse (CDW). The demographic variables included race, age, sex, body mass index, military service period, combat service, exposure to Agent Orange, exposure to ionizing radiation, prisoner of war (POW) status, unemployability status, housebound status, military service-connected disability, Medicaid eligibility, and receipt of VA pension or social security benefits. Ever-smokers were identified based on ICD9/10CM diagnosis codes and health factor indicators extracted from electronic health records.

The inpatient and outpatient claims data were utilized to identify the year of lung cancer diagnosis and determine if veterans received treatment for 16 chronic conditions. These conditions encompassed anemia, congestive heart failure, coronary artery disease, dementia, diabetes, obesity, hypertension, hypothyroidism,

[8]Other choices for the match restrictions are possible that increase the match rate but decrease the accuracy. Spot-checking 1000 random matched addresses found 4 mismatches under the chosen algorithm. Note that the algorithm requires perfect matches on the property number, city, and zip code, so mismatches most likely occur within a reasonable distance to the true location

rheumatoid arthritis, pulmonary fibrosis, asthma, chronic obstructive pulmonary disorder (COPD), drug/alcohol dependence, drug/alcohol abuse, and alcohol use.

Furthermore, estimates of atmospheric bismuth-214 and indoor radon exposure were generated using the models developed in Watkins (2024) for each month that a veteran resided in the matched property. These monthly estimates were averaged over each year to create a measure of mean long-term radon exposure. Cumulative sums of these estimates were calculated for each year that a veteran was observed in the panel.

Table 1 presents summary statistics for the unique panel of veterans, while the summary statistics of the yearly panel of observations are provided in Table 2. It is worth noting that the average indoor radon prediction for this population is relatively low, with an average of approximately 1.6 piC/L. It is important to emphasize that this value represents the yearly average of short-term predictions. While multiplying these yearly means by 365 would estimate the total yearly exposure, it is common practice in the literature and EPA guidelines to present thresholds and findings based on yearly means. On average, veterans are observed for a period of 5.4 years, accumulating an average cumulative exposure of 8.7 piC/L.

## 0.4 Empirical Strategy

### 0.4.1 Lung Cancer

The objective of this paper is to estimate the health effects of cumulative radon exposure among US veterans. I model the relationship between lung cancer and cumulative radon exposure using the following regression equation:

$$LungCancer_{ict} = \beta log(r_{ict}) + X'_{ict}\gamma + \alpha_c + \alpha_t + \alpha_{ct} + c_i + e_{ict} \tag{2}$$

where $LungCancer_{ict}$ is an indicator of lung cancer for individual $i$, who lives in county $c$, in year $t$. The coefficient of interest is $\beta$, the marginal effect of cumulative indoor radon exposure ($r_{ict}$). $X'_{ict}$ is a vector of veteran characteristics (smoking status, race, Medicaid eligibility, flags for receiving VA disability receiving social security benefits, if the veteran is considered unemployable, if the veteran is housebound, history of agent orange and/or ionizing radiation exposure, combat service, military service period, sex, and second-order polynomials of BMI and age, yearly averages of precipitation, minimum temperature, maximum temperature, dew point temperature, and vapor pressure deficits).

In order to account for various sources of heterogeneity and control for unobserved factors, this study incorporates a high dimension of fixed effects. Specifically, county fixed effects ($\alpha_c$), year fixed effects ($\alpha_t$), and county-by-year fixed effects ($\alpha_{ct}$) are included in the analysis. The county fixed effects help control for constant geographic differences that may impact lung cancer rates. The year fixed effects account for yearly shocks that could affect the development or diagnosis of lung cancer among veterans as a whole. Lastly, the

Table 1: Veteran Cohort Summary Statistics

| | Mean | Std Dev | N |
|---|---|---|---|
| Lung Cancer | 0.027 | 0.162 | 700,255 |
| Died | 0.375 | 0.484 | 700,255 |
| Hypertension | 0.551 | 0.497 | 700,255 |
| Coronary Artery Disease | 0.237 | 0.425 | 700,255 |
| Anemia | 0.204 | 0.403 | 700,255 |
| Pulmonary Fibrosis | 0.005 | 0.073 | 700,255 |
| COPD | 0.167 | 0.373 | 700,255 |
| Rheumatoid Arthritis | 0.027 | 0.162 | 700,255 |
| Obesity | 0.139 | 0.346 | 700,255 |
| Hypothyroidism | 0.085 | 0.279 | 700,255 |
| Drug/Alcohol Use | 0.008 | 0.086 | 700,255 |
| Drug/Alcohol Dependence | 0.057 | 0.233 | 700,255 |
| Drug Abuse | 0.168 | 0.374 | 700,255 |
| Diabetes | 0.275 | 0.446 | 700,255 |
| Dementia | 0.060 | 0.237 | 700,255 |
| Congestive Heart Failure | 0.080 | 0.272 | 700,255 |
| Asthma | 0.037 | 0.190 | 700,255 |
| Race - White | 0.806 | 0.395 | 700,255 |
| Race - Black | 0.179 | 0.383 | 700,255 |
| Race - Other | 0.015 | 0.122 | 700,255 |
| Medicaid Eligible | 0.009 | 0.084 | 700,255 |
| Service Connected Disability | 0.537 | 0.499 | 700,255 |
| Agent Orange Exposure | 0.088 | 0.283 | 700,255 |
| Ionizing Radiation Exposure | 0.002 | 0.046 | 700,255 |
| SW Asia Service | 0.024 | 0.150 | 700,255 |
| Prisoner of War | 0.004 | 0.048 | 700,255 |
| Combat Service | 0.171 | 0.350 | 700,255 |
| Unemployable | 0.066 | 0.248 | 700,255 |
| VA Disability | 0.483 | 0.499 | 700,255 |
| PT Disability | 0.182 | 0.381 | 700,255 |
| Housebound | 0.044 | 0.205 | 700,255 |
| Receiving VA Pension | 0.050 | 0.214 | 700,255 |
| Receiving SS Benefit | 0.131 | 0.258 | 700,255 |
| Service Period - Korean | 0.092 | 0.289 | 700,255 |
| Service Period - Vietnam | 0.354 | 0.478 | 700,255 |
| Service Period - WWII | 0.102 | 0.303 | 700,255 |
| Service Period - Other | 0.452 | 0.498 | 700,255 |
| Ever-Smoker | 0.373 | 0.483 | 700,255 |
| BMI | 29.230 | 5.307 | 700,255 |

Notes: Table reports summary statistics for the unique cohort of veterans.

Table 2: Veteran Yearly Panel Summary Statistics

| | Mean | Std Dev | N |
|---|---|---|---|
| Lung Cancer | 0.010 | 0.098 | 4,470,821 |
| Died | 0.059 | 0.235 | 4,470,821 |
| Year | 2014 | 4.9 | 4,470,821 |
| Years in Panel | 5.406 | 4.065 | 4,470,821 |
| Yearly Mean Radon piC/L | 1.625 | 0.656 | 4,470,821 |
| Yealy Mean Atmospheric Bi214 counts/sec | 0.00005 | 0.00003 | 4,470,821 |
| Cumulative Radon piC/L | 8.736 | 7.835 | 4,470,821 |
| Cumulative Bi214 counts/sec | 0.0003 | 0.0003 | 4,470,821 |
| Precipitaiton | 4.033 | 0.934 | 4,470,821 |
| Maximum Temperature | 22.627 | 2.223 | 4,470,821 |
| Minimum Temperature | 10.934 | 2.483 | 4,470,821 |
| Mean Dew Point Temperature | 10.383 | 2.564 | 4,470,821 |
| Maximum Vapor Pressure Deficit | 15.731 | 2.300 | 4,470,821 |
| Minimum Vapor Pressure Deficit | 1.005 | 0.534 | 4,470,821 |

Notes: Table reports summary statistics for the yearly panel dataset of the cohort of veterans.

county-by-year fixed effects flexibly capture location-time-varying shocks, such as changes in VHA policies or other local pollutants.

In addition to these fixed effects, individual-specific unobserved heterogeneity ($c_i$) is considered in the model. Accounting for $c_i$ is particularly crucial in this research context as it addresses the potential correlation between unobserved individual characteristics and predicted radon exposure. If the unobserved heterogeneity ($c_i$) is uncorrelated with the predicted long-term radon ($\hat{r_{it}}$) and other confounding factors (such as agent orange exposure, smoking status, race, etc.), pooled OLS and random effects (RE) estimators provide consistent estimates of $\beta$ (with random effects being more efficient). However, if $c_i$ includes previous radon exposure and there is systematic correlation with $\hat{r_{it}}$, pooled OLS and RE estimators would produce inconsistent results. Similarly, if a veteran's smoking history is correlated with their childhood second-hand smoke exposure and an interaction term is estimated between $\hat{r_{it}}$ and smoking status, the coefficients on these interaction terms would be inconsistent without controlling for $c_i$.

Furthermore, controlling for $c_i$ is essential in any analysis involving cumulative exposure. The average number of years observed in the panel is 5.4, which is insufficient to capture the full latency period for developing lung cancer, ranging from 10-30 years (Lipfert and Wyzga, 2019), and potentially up to 55 years (Hu et al., 1994). Given these timeframes and the low levels of cumulative predicted exposure among this population, it would not be expected to observe a significant positive coefficient on cumulative radon and lung cancer risk. In fact, one might even anticipate a negative coefficient for low levels of cumulative exposure due to the use of low radiation levels in radiation therapy. Indeed, Scott (2011) finds that low levels of radon

exposure inhibit the development of lung cancer. To control for linear unobserved heterogeneity, the fixed-effects (within) transformation is applied, which subtracts the means of the variables for each veteran. This approach helps address the potential biases arising from unobserved time-invariant factors and allows for more reliable estimation of the relationship between cumulative radon exposure and lung cancer risk.:

$$\Delta LungCancer_{ict} = \beta \Delta log(r_{ict}) + \Delta X'_{ict}\gamma + \Delta\alpha_t + \Delta\alpha_{ct} + \Delta e_{ict} \tag{3}$$

where $\Delta r_{itsyg}$ represents $r_{itsyg} = r_{itsyg} - \bar{r}_{itsyg}$ under the fixed-effect (within) transformation. This model relies on the crucial assumption that $c_i$ does not interact with observable characteristics and is constant over time. I leave the analysis that allows $c_i$ to vary over time or interact with the observable characteristics, such as random coefficient or correlated random effect models, as future work.

To address the limitations associated with distributional assumptions and incidental parameters in a panel probit or logit model within an individual fixed effects framework, a linear probability model (LPM) is employed in this study. By utilizing the LPM as a linear projection, the analysis aims to examine the significance of long-term radon exposure on the probability of developing lung cancer. The LPM provides straightforward interpretations and facilitates the estimation of marginal effects.

Furthermore, to account for the sampling uncertainty of the generated regressor $\hat{r_{it}}$, nonparametric bootstrapped coefficient and standard error estimates are presented in all results. Although the panel consists of 4.5 million observations among 700K unique veterans, the sampling uncertainty is likely to be relatively small. To address this uncertainty, the unique list of veterans is bootstrapped 100 times, with all yearly observations for the sampled individuals retained. In cases where a veteran is sampled multiple times, each set of observations is treated as a separate individual in the analysis. It is important to note that increasing the number of bootstraps significantly escalates computational costs while providing minimal gains in precision due to the large sample sizes. Hence, the decision to conduct 100 bootstraps strikes a balance between computational efficiency and obtaining reliable estimates for the regression coefficients and their standard errors.

### 0.4.2 Additional Models

#### 0.4.2.1 Atmospheric Bi-214

I include an analysis of atmospheric bi-214 and lung cancer risk:

$$\Delta LungCancer_{ict} = \gamma \Delta log(Bi214_{ict}) + \Delta X'_{ict}\gamma + \Delta\alpha_t + \Delta\alpha_{ct} + \Delta e_{ict} \tag{4}$$

#### 0.4.2.2 Death

I model the effect of long-term radon exposure on death using the following reduced-form equation, also utilizing the fixed-effects/within transformation:

$$\Delta Death_{ict} = \pi_1 \Delta log(r_{ict}) + \Delta X'_{ict}\pi_2 + \Delta Z'_{ict}\pi_3 + \Delta\alpha_t + \Delta\alpha_{ct} + \Delta u_{ict} \tag{5}$$

where $r_{ict}$, $X'_{ict}$, and $\alpha$'s are the same as above, $Z'_{ict}$ is a vector of the chronic conditions (lung and other severe cancers, hypertension, coronary artery disease, anemia, pulmonary fibrosis, COPD, rheumatoid arthritis, obesity, hypothyroidism, drub abuse/dependence/use, diabetes, dementia, heart failure), and $u_{ict}$ is the error term.

#### 0.4.2.3 Relative Risks and Radon Attributed Fraction

This paper differs from the literature composed of case-control studies and meta-analyses in the US by providing the first analysis of a panel of individuals over time. The results are in the form of marginal effects and probabilities of lung cancer and death, whereas the epidemiological literature provides odds ratios and relative risks. However, since the models are linear, it is relatively straightforward to generate results that are comparable to the epidemiological literature on indoor radon exposure. Relative risks can be defined as:

$$RR = \frac{p(LungCancer = 1|R = r+1, X)}{p(LungCancer = 1|R = r, X)} \tag{6}$$

where the numerator of equation 6 represents the probability of developing lung cancer from increasing cumulative radon exposure by 1 piC/L. The denominator represents the baseline probability of lung cancer risk among this population. I use the regression results to calculate estimates of the RR at the mean. Observe that the denominator is approximately equal to the dependent variable mean. Then, since in the specification that uses the natural log of cumulative radon exposure, the $p(LungCancer = 1|R = r+1, X)$ is approximately equal to $\beta log[(r+1)/r]$ at the mean. Within each bootstrap sample, I estimate the following as a measure of relative risks:

$$\hat{RR} \approx \frac{\hat{\beta}\ log[(\bar{r}+1)/\bar{r}]}{\bar{y}_{lung}}. \tag{7}$$

where $\bar{y}_{lung}$ is the mean probability of lung cancer.

In population terms, the radon-attributed fraction (RAF) is defined as the number of lung cancer deaths attributed to radon exposure. I use a similar approach to estimate the RAF using the following equation:

$$RAF = \frac{p(LungCancer = 1|R = r+1, X)}{p(Died = 1|R = r+1, X)} \tag{8}$$

where the numerator is the same as in equation 6, and the denominator represents the probability of dying from increasing radon exposure by one unit. To include estimates of the RAF, consider the following. The total number of observed lung cancer cases is approximately equal to the predicted probability of lung cancer multiplied by the total number of observations (N). It follows that the increase in lung cancer cases from increasing cumulative radon is simply the marginal effect multiplied by N. Similarly, for deaths attributed to radon exposure: the total increase in deaths attributed to radon exposure is simply the marginal effect multiplied by N. Thus, one can approximate $p(LungCancer = 1|R = r+1, X) \approx \bar{y}_{lung} + \hat{\beta} \cdot log[(\bar{r}+1)/\bar{r}]$ when using the log specification of cumulative radon. Similarly, one can approximate $p(Died = 1|R = r+1, X) \approx \bar{y}_{died} + \hat{\pi}_1 \cdot log[(\bar{r}+1)/\bar{r}]$, where $\bar{y}_{died}$ is the mean probability of death. Within each bootstrapped sample, I estimate the following:

$$\hat{RAF} \approx \frac{\bar{y}_{lung} + \hat{\beta} \cdot log[(\bar{r}+1)/\bar{r}]}{\bar{y}_{died} + \hat{\pi}_1 \cdot log[(\bar{r}+1)/\bar{r}]}. \tag{9}$$

### 0.5 Results

The analysis initially focuses on the impact of EPA Zone Designations, represented by three indicator variables, on the probability of lung cancer. Table 3 presents these results, with EPA Zone 3 as the reference category, and each column representing a different combination of State by year fixed effects. It should be noted that the EPA Zone is a county-level measure, and including county-fixed effects eliminates the possibility of estimating the parameter for EPA Zone. Similarly, employing individual fixed effects through a within-transformation also removes the EPA measure from the regression. While these results do not establish causal relationships, they provide valuable descriptive insights.

Compared to the mean lung cancer risk, veterans residing in EPA Zone 1 counties are approximately 12.5 percent more likely to receive a lung cancer diagnosis, relative to veterans in EPA Zone 3 counties. No statistically significant difference in lung cancer risk is observed between Zones 2 and 3. Nevertheless, there are several limitations when utilizing the EPA measure. Unobservable characteristics correlated with the EPA zone designation at the county level cannot be controlled for in the analysis. Furthermore, interpreting the results from a public policy perspective poses challenges, as EPA Zone 1 is already designated as a high-risk zone, and the parameter estimates largely confirm the correlation between higher radon levels and lung cancer risk among veterans. The estimates do not provide guidance on appropriate mitigation levels or acceptable exposure thresholds.

Additional issues related to sorting and avoidance behavior may complicate the analysis, underscoring the importance of leveraging exogenous changes in radon risks. Access to data before and after the release of the EPA map in 1993 could potentially enable the identification of reduced-form effects of indoor radon

Table 3: EPA Zone on Lung Cancer Probability among Veterans

| | Probability of Lung Cancer | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| EPA Zone 1 | 0.00120 | 0.00123 | 0.00117 |
| | (0.00028, 0.00190) | (0.00034, 0.00194) | (0.00030, 0.00190) |
| EPA Zone 2 | 0.00048 | 0.00050 | 0.00045 |
| | (−0.00019, 0.00121) | (−0.00015, 0.00123) | (−0.00023, 0.00117) |
| Mean Dependent Variable | 0.0096 | 0.0096 | 0.0096 |
| Radon Mean | 8.74 | 8.74 | 8.74 |
| Zone 1 Effect Relative to Mean | 12.5% | 12.8% | 12.2% |
| Zone 2 Effect Relative to Mean | 5% | 5.3% | 4.7% |
| **Fixed Effects** | | | |
| Veteran | No | No | No |
| State by Year | Yes | No | No |
| State by 2-Year | No | Yes | No |
| State by 3-Year | No | No | Yes |

Notes: Table reports estimates of 2, using the EPA Zone designation as the measure of indoor radon. The excluded category is the EPA Zone 3, the lowest risk designation. Each column includes a different set of State and year fixed effects, since county fixed effects absorb the variation in the EPA Zone. All regressions include the veteran covariates of smoking status, race, Medicaid eligibility, flag for receiving VA disability and/or receiving social security benefits, agent orange exposure, combat service, sex, second-order polynomials of BMI and age, yearly weather (precipitation, min and max temperatures, min and max dew point temperatures, vapor pressure deficit). Bootstrapped 95 precent confidence intervals are given in parenthesis.

exposure on the probability of lung cancer. However, it is essential to recognize that estimates based on perceived indoor radon risk alone cannot ascertain the precise levels of radon exposure that cause lung cancer or capture the marginal effects of continuous exposure. This is where the structural estimates of radon exposure on lung cancer risk, as offered by the models in Chapter 1, provide a notable contribution.

The analysis proceeds by estimating equations 3 and 4, incorporating individual veteran fixed effects (equivalent to the within-transformation). The results are presented in Table 4 The parameter estimates for cumulative atmospheric bi-214 and cumulative indoor radon are remarkably similar: a 10 percent increase in either measure corresponds to a 6.2 percent increase in the probability of lung cancer. These findings hold across various specifications and demonstrate both economic and statistical significance at the 1 percent level. The striking similarity of the estimates obtained from different measures, namely atmospheric bi-214 and indoor radon, provides reassurance regarding the accurate representation of atmospheric radon-222 by atmospheric bi-214.

The magnitude of these effects is considerable, particularly when considering that the average yearly indoor radon concentration is 1.6 piC/L, resulting in an average cumulative exposure of 8.7 piC/L. Importantly, these effects persist even at levels of exposure **below** the suggested remediation threshold of 4 piC/L. The results underscore the significant lung cancer risk posed by radon exposure among the veteran population examined in this study.

Table 4: Bi-214 and Radon on Lung Cancer Probability among Veterans

| | Probability of Lung Cancer | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| ***Panel A: Cumulative Bi-214*** | | | |
| Log Cumulative Bi214 | 0.00622 | 0.00646 | 0.00657 |
| | (0.00602, 0.00643) | (0.00627, 0.00668) | (0.00636, 0.00679) |
| Mean Dependent Variable | 0.0096 | 0.0096 | 0.0096 |
| Bi-214 Mean | 0.00026 | 0.00026 | 0.00026 |
| 10% Effect Relative to Mean | 6.2% | 6.2% | 6.2% |
| Relative Risk | 1.07 | 1.07 | 1.07 |
| ***Panel B: Cumulative Indoor Radon*** | | | |
| Log Cumulative Radon | 0.00624 | 0.00648 | 0.00658 |
| | (0.00603, 0.00646) | (0.00629, 0.00671) | (0.00637, 0.00681) |
| Mean Dependent Variable | 0.0096 | 0.0096 | 0.0096 |
| Radon Mean | 8.74 | 8.74 | 8.74 |
| 10% Effect Relative to Mean | 6.2% | 6.2% | 6.2% |
| Relative Risk | 1.07 | 1.07 | 1.07 |
| **Fixed Effects** | | | |
| Veteran | Yes | Yes | Yes |
| County by Year | Yes | No | No |
| County by 2-Year | No | Yes | No |
| County by 3-Year | No | No | Yes |

Notes: Table reports fixed-effect (within) estimates of equations 3 and 4. Panel A uses the natural log of cumulative atmospheric bismuth-214 as the main variable of interest. Panel B uses the natural log of cumulative predicted indoor radon. Each column includes a different set of County and year fixed effects. All regressions include the veteran covariates of smoking status, race, Medicaid eligibility, flag for receiving VA disability and/or receiving social security benefits, agent orange exposure, combat service, sex, second-order polynomials of BMI and age, yearly weather (precipitation, min and max temperatures, min and max dew point temperatures, vapor pressure deficit). Bootstrapped 95 precent confidence intervals are given in parenthesis.

Furthermore, the analysis explores the heterogeneous effects of radon exposure. The corresponding tables are provided in the Appendix and consistently support a coherent narrative. Veterans with higher suscepti-

bility to lung disease, such as ever-smokers and those diagnosed with COPD, exhibit substantially higher marginal effects of cumulative radon exposure on lung cancer risk. However, it is noteworthy that the relative effects compared to the mean cancer risk are relatively consistent across different populations. For instance, Table 6 presents the results after stratifying the analysis for ever-smokers and never-smokers. In both cases, the effects of radon on lung cancer risk are approximately 6 percent relative to the mean risk. A similar pattern emerges when comparing veterans with a COPD diagnosis to those without. Conversely, non-combat veterans display slightly higher marginal effects relative to their mean lung cancer risk compared to combat veterans, albeit only marginally larger (5.1 percent vs. 6.3 percent). In summary, while the marginal effects of radon exposure can vary significantly across different populations, they remain relatively consistent when considering the **relative risks** in relation to the baseline risk for each respective population.

Moving forward, I present results from estimating the radon-attributed fraction (RAF) using Equation 9. Particularly interesting are the comparisons between ever-smokers and never-smokers, as well as between veterans with a COPD diagnosis and those without. Table 5 displays the results for ever-smokers vs. never-smokers. The estimated RAF for ever-smoking veterans is approximately 28 percent, while the RAF for never-smokers is only 8 percent. These findings are significantly different and somewhat unexpected. One might anticipate that the RAF would be higher for the never-smoking population, considering that their primary risk factor for lung cancer is radon exposure. However, a plausible explanation for this result is that the level of indoor radon exposure is relatively low, and therefore does not generate sufficient risk to cause radon-attributed lung cancer deaths among the never-smokers. It implies that non-smokers are not experiencing lung cancer deaths at proportions similar to the smoking population, which aligns with the understanding that radon primarily contributes to death through its impact on comorbidities. A similar pattern emerges for veterans with COPD: the RAF for veterans diagnosed with COPD is approximately 35 percent, whereas for those without COPD it is only 8 percent. These results raise an important concern: populations with a significant risk of lung cancer death are at significantly higher risk of radon exposure, even at relatively low levels. However, it should be noted that these results assume that the models accurately capture the causal effects at the mean in both the lung cancer and death specifications.

Nevertheless, the evidence strongly suggests that veterans with COPD and smoking histories face an elevated risk of indoor radon exposure. Therefore, it is crucial for the VHA to take action by notifying these veterans and developing appropriate plans to reduce their exposure. Overall, the differential marginal effects of radon exposure among various subpopulations within the VA underscore the importance of tailored public health policies, targeted interventions, prevention strategies, awareness campaigns, and research endeavors. By addressing the specific risks and vulnerabilities of these populations, such as smokers, non-smokers, and individuals with COPD, health policies can effectively alleviate the burden of lung cancer associated with

Table 5: Radon-Attributed Fraction by Subpopulation

| | Radon-Attributed Fraction | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| ***Panel A: Ever-Smoking Veterans*** | | | |
| Radon-Attributed Fraction | 0.282 | 0.282 | 0.282 |
| | (0.273, 0.290) | (0.273, 0.290) | (0.273, 0.291) |
| Mean Lung Cancer | 0.0186 | 0.0186 | 0.0186 |
| Mean Death | 0.0615 | 0.0615 | 0.0615 |
| Radon Mean | 8.34 | 8.34 | 8.34 |
| ***Panel B: Never-Smoking Veterans*** | | | |
| Radon-Attributed Fraction | 0.078 | 0.078 | 0.078 |
| | (0.074, 0.081) | (0.074, 0.081) | (0.074, 0.081) |
| Mean Lung Cancer | 0.0048 | 0.0048 | 0.0048 |
| Mean Death | 0.0573 | 0.0573 | 0.0573 |
| Radon Mean | 8.95 | 8.95 | 8.95 |
| ***Panel C: Veterans with COPD*** | | | |
| Radon-Attributed Fraction | 0.356 | 0.355 | 0.355 |
| | (0.347, 0.365) | (0.346, 0.364) | (0.346, 0.364) |
| Mean Lung Cancer | 0.041 | 0.041 | 0.041 |
| Mean Death | 0.1056 | 0.1056 | 0.1056 |
| Radon Mean | 7.73 | 7.73 | 7.73 |
| ***Panel D: Veterans without COPD*** | | | |
| Radon-Attributed Fraction | 0.078 | 0.078 | 0.078 |
| | (0.075, 0.081) | (0.075, 0.081) | (0.075, 0.081) |
| Mean Lung Cancer | 0.0043 | 0.0043 | 0.0043 |
| Mean Death | 0.0507 | 0.0507 | 0.0507 |
| Radon Mean | 8.91 | 8.91 | 8.91 |
| **Fixed Effects** | | | |
| Veteran | Yes | Yes | Yes |
| County by Year | Yes | No | No |
| County by 2-Year | No | Yes | No |
| County by 3-Year | No | No | Yes |

Notes: Table reports estimates of 9 using the parameter estimates from the fixed-effect (within) results of radon on lung cancer and radon on death. Each panel presents the results among a different set of subpopulations. Panel A subsets the data to only those veterans with a history of smoking. Panel B uses only those veterans with no history of smoking. Panel C uses veterans with a COPD diagnosis. Panel B subsets the data to only those veterans with no COPD diagnosis. Each column includes a different set of County and year fixed effects. All regressions include the veteran covariates of smoking status, race, Medicaid eligibility, flag for receiving VA disability and/or receiving social security benefits, agent orange exposure, copd service, sex, second-order polynomials of BMI and age, yearly weather (precipitation, min and max temperatures, min and max dew point temperatures, vapor pressure deficit). Bootstrapped 95 percent confidence errors are given in parenthesis.

radon exposure.

### 0.6 Conclusion

This work has significantly improved our understanding of the health effects of indoor radon exposure among United States veterans. By leveraging predictive models and unique datasets, I am able to estimate the effects of long-term cumulative radon exposure on the probability of developing lung cancer, providing valuable insights into the risks faced by this specific population.

The findings highlight the substantial impact of radon exposure on the probability of developing lung cancer among veterans, even at relatively low levels of exposure. A significant association between cumulative radon exposure and the increased likelihood of lung cancer, particularly among subpopulations at higher risk, such as ever-smokers and veterans diagnosed with chronic obstructive pulmonary disease (COPD). These results emphasize the importance of targeted interventions and preventive strategies to reduce the burden of lung cancer in these vulnerable groups.

Furthermore, the analysis of the radon-attributed fraction (RAF) revealed the proportion of lung cancer deaths that can be attributed to radon exposure within specific veteran subpopulations. Veterans with COPD exhibited a particularly high risk, with a substantial fraction of lung cancer deaths (35 percent) attributable to radon. Similarly, ever-smokers had a RAF of approximately 28 percent. These findings underscore the urgent need for tailored public health policies and interventions aimed at reducing radon exposure and its associated health risks among high-risk individuals.

The comprehensive nature of this study allowed us to control for important confounding factors and individual heterogeneity through fixed-effects models. This approach strengthens the validity of the findings and provides a robust estimation of the effects of radon exposure on lung cancer risk within the veteran population. However, the linear assumptions are critical in the interpretation of the results.

While this study has shed light on the health risks associated with radon exposure, further research is needed to refine estimates and deepen our understanding of the mechanisms underlying these effects. In particular, future work should investigate potential nonlinear models that allow for interactions between unobserved heterogeneity or utilize address histories to more accurately capture true cumulative radon exposure. Collaboration between the Veterans Health Administration (VHA) and the Environmental Protection Agency (EPA) could facilitate ongoing efforts to develop more effective mitigation strategies, establish appropriate radon exposure thresholds specifically tailored for veterans, and promote awareness of the risks posed by radon.

In conclusion, this study provides compelling evidence of the detrimental effects of indoor radon exposure on the probability of developing lung cancer among United States veterans. The findings emphasize the need

for targeted interventions, awareness campaigns, and public health policies aimed at reducing radon exposure and mitigating the burden of lung cancer within the veteran population. By addressing the specific risks faced by different subpopulations, we can work towards minimizing health disparities associated with radon exposure and improving the overall well-being of veterans.

## 0.7 Appendix

### 0.7.1 Additional Results

Table 6: Radon on Lung Cancer Probability by Smoking History

| | Probability of Lung Cancer | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| ***Panel A: Ever-smoking Veterans*** | | | |
| Log Cumulative Radon | 0.01122 | 0.01158 | 0.01177 |
| | (0.01069, 0.01189) | (0.01105, 0.01224) | (0.01124, 0.01247) |
| Mean Dependent Variable | 0.0186 | 0.0186 | 0.0186 |
| Radon Mean | 8.34 | 8.34 | 8.34 |
| 10% Effect Relative to Mean | 5.7% | 5.7% | 5.7% |
| Relative Risk | 1.068 | 1.068 | 1.068 |
| ***Panel B: Never-Smoking Veterans*** | | | |
| Log Cumulative Radon | 0.00277 | 0.00300 | 0.00304 |
| | (0.00253, 0.00295) | (0.00275, 0.00320) | (0.00278, 0.00323) |
| Mean Dependent Variable | 0.0048 | 0.0048 | 0.0048 |
| Radon Mean | 8.95 | 8.95 | 8.95 |
| 10% Effect Relative to Mean | 5.5% | 5.5% | 5.5% |
| Relative Risk | 1.061 | 1.061 | 1.061 |
| **Fixed Effects** | | | |
| Veteran | Yes | Yes | Yes |
| County by Year | Yes | No | No |
| County by 2-Year | No | Yes | No |
| County by 3-Year | No | No | Yes |

Notes: Table reports estimates of 3 using the fixed-effect (within) estimator. Panel A subsets the data to only those veterans with a history of smoking. Panel B subsets the data to only those veterans with no history of smoking Each column includes a different set of County and year fixed effects. All regressions include the veteran covariates of smoking status, race, Medicaid eligibility, flag for receiving VA disability and/or receiving social security benefits, agent orange exposure, combat service, sex, second-order polynomials of BMI and age, yearly weather (precipitation, min and max temperatures, min and max dew point temperatures, vapor pressure deficit). Bootstrapped 95 precent confidence intervals are given in parenthesis.

Table 7: Radon on Lung Cancer Probability by COPD Diagnosis

| | Probability of Lung Cancer | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| ***Panel A: Veterans with COPD*** | | | |
| Log Cumulative Radon | 0.02153 | 0.02252 | 0.02298 |
| | (0.01939, 0.02279) | (0.02048, 0.02355) | (0.02107, 0.02415) |
| Mean Dependent Variable | 0.041 | 0.041 | 0.041 |
| Radon Mean | 7.73 | 7.73 | 7.73 |
| 10% Effect Relative to Mean | 5% | 5% | 5% |
| Relative Risk | 1.064 | 1.064 | 1.064 |
| ***Panel B: Veterans without COPD*** | | | |
| Log Cumulative Radon | 0.00257 | 0.00269 | 0.00274 |
| | (0.00240, 0.00272) | (0.00253, 0.00284) | (0.00257, 0.00289) |
| Mean Dependent Variable | 0.0043 | 0.0043 | 0.0043 |
| Radon Mean | 8.91 | 8.91 | 8.91 |
| 10% Effect Relative to Mean | 5.8% | 5.8% | 5.8% |
| Relative Risk | 1.064 | 1.064 | 1.064 |
| **Fixed Effects** | | | |
| Veteran | Yes | Yes | Yes |
| County by Year | Yes | No | No |
| County by 2-Year | No | Yes | No |
| County by 3-Year | No | No | Yes |

Notes: Table reports estimates of 3 using the fixed-effect (within) estimator. Panel A subsets the data to only those veterans that have a diagnosis of COPD. Panel B subsets the data to only those veterans without a COPD diagnosis. Each column includes a different set of County and year fixed effects. All regressions include the veteran covariates of smoking status, race, Medicaid eligibility, flag for receiving VA disability and/or receiving social security benefits, agent orange exposure, combat service, sex, second-order polynomials of BMI and age, yearly weather (precipitation, min and max temperatures, min and max dew point temperatures, vapor pressure deficit). Bootstrapped 95 precent confidence intervals are given in parenthesis.

Table 8: Radon on Lung Cancer Probability by Combat History

| | Probability of Lung Cancer | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| ***Panel A: Combat Veterans*** | | | |
| Log Cumulative Radon | 0.00635 | 0.00692 | 0.00719 |
| | (0.00553, 0.00694) | (0.00614, 0.00744) | (0.00638, 0.00781) |
| Mean Dependent Variable | 0.0119 | 0.0119 | 0.0119 |
| Radon Mean | 7.87 | 7.87 | 7.87 |
| 10% Effect Relative to Mean | 5.1% | 5.1% | 5.1% |
| Relative Risk | 1.064 | 1.064 | 1.064 |
| ***Panel B: Non-Combat Veterans*** | | | |
| Log Cumulative Radon | 0.00613 | 0.00633 | 0.00643 |
| | (0.00584, 0.00644) | (0.00604, 0.00665) | (0.00615, 0.00675) |
| Mean Dependent Variable | 0.0092 | 0.0092 | 0.0092 |
| Radon Mean | 8.87 | 8.87 | 8.87 |
| 10% Effect Relative to Mean | 6.3% | 6.3% | 6.3% |
| Relative Risk | 1.071 | 1.071 | 1.071 |
| **Fixed Effects** | | | |
| Veteran | Yes | Yes | Yes |
| County by Year | Yes | No | No |
| County by 2-Year | No | Yes | No |
| County by 3-Year | No | No | Yes |

Notes: Table reports estimates of 3 using the fixed-effect (within) estimator. Panel A subsets the data to only those veterans with a history of combat Panel B subsets the data to only those veterans with no history of combat. Each column includes a different set of County and year fixed effects. All regressions include the veteran covariates of smoking status, race, Medicaid eligibility, flag for receiving VA disability and/or receiving social security benefits, agent orange exposure, combat service, sex, second-order polynomials of BMI and age, yearly weather (precipitation, min and max temperatures, min and max dew point temperatures, vapor pressure deficit). Bootstrapped 95 precent confidence intervals are given in parenthesis.

Table 9: Radon-Attributed Fraction by Combat History

| | Radon-Attributed Fraction | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| ***Panel A: Combat Veterans*** | | | |
| Radon-Attributed Fraction | 0.140 | 0.139 | 0.140 |
| | (0.134, 0.149) | (0.134, 0.149) | (0.134, 0.149) |
| Mean Lung Cancer | 0.0119 | 0.0119 | 0.0119 |
| Mean Death | 0.0781 | 0.0781 | 0.0781 |
| Radon Mean | 7.87 | 7.87 | 7.87 |
| ***Panel B: Non-Combat Veterans*** | | | |
| Radon-Attributed Fraction | 0.155 | 0.155 | 0.155 |
| | (0.152, 0.159) | (0.152, 0.159) | (0.152, 0.159) |
| Mean Lung Cancer | 0.0092 | 0.0092 | 0.0092 |
| Mean Death | 0.0558 | 0.0558 | 0.0558 |
| Radon Mean | 8.87 | 8.87 | 8.87 |
| **Fixed Effects** | | | |
| Veteran | Yes | Yes | Yes |
| County by Year | Yes | No | No |
| County by 2-Year | No | Yes | No |
| County by 3-Year | No | No | Yes |

Notes: Table reports estimates of 9 using the parameter estimates from the fixed-effect (within) results of radon on lung cancer and radon on death. Panel A subsets the data to only those veterans with a history of combat. Panel B subsets the data to only those veterans with no history of combat. Each column includes a different set of County and year fixed effects. All regressions include the veteran covariates of smoking status, race, Medicaid eligibility, flag for receiving VA disability and/or receiving social security benefits, agent orange exposure, combat service, sex, second-order polynomials of BMI and age, yearly weather (precipitation, min and max temperatures, min and max dew point temperatures, vapor pressure deficit). Bootstrapped standard errors are given in parenthesis.

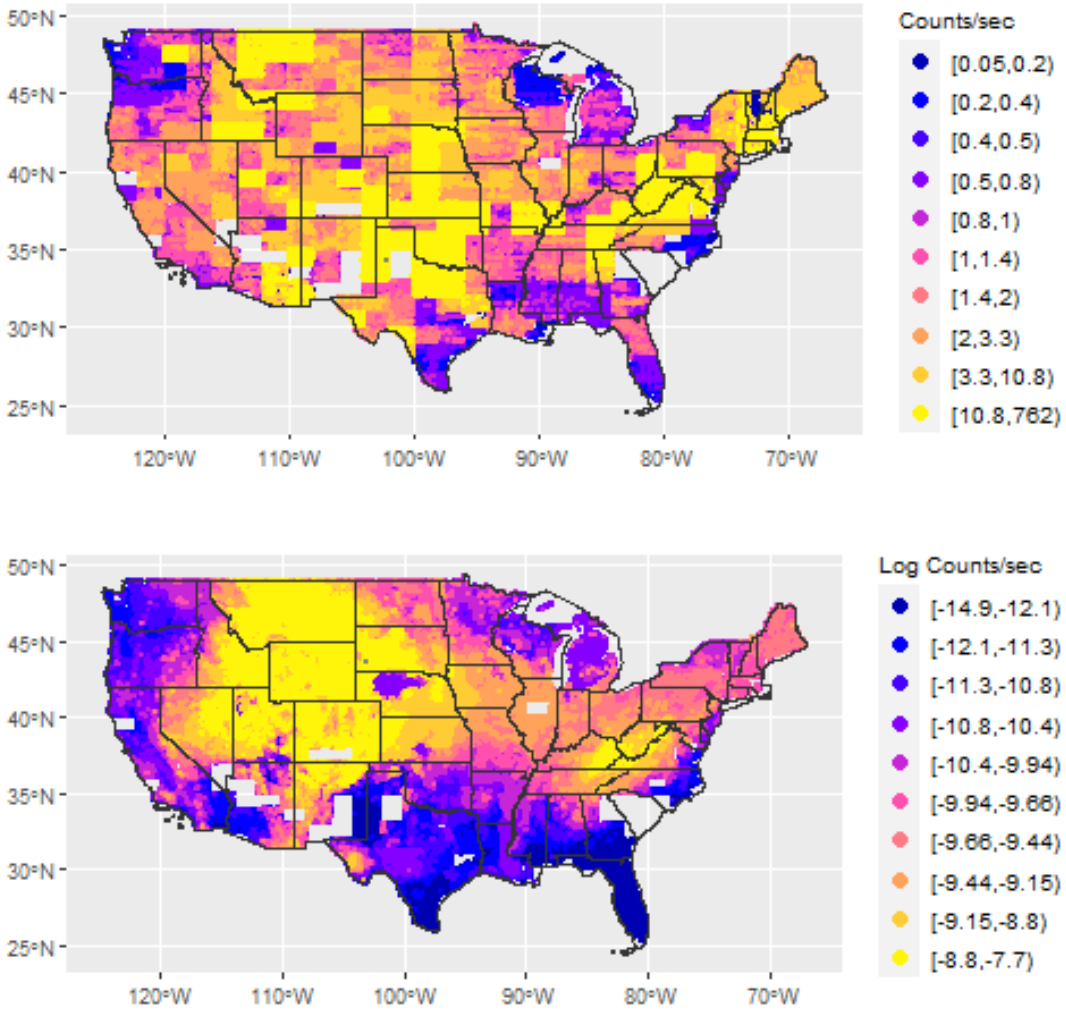
Table 10: Radon on Lung Cancer Probability by EPA Zone

| | Probability of Lung Cancer | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| ***Panel A: Veterans in EPA Zone 1*** | | | |
| Log Cumulative Radon | 0.00715 | 0.00743 | 0.00760 |
| | (0.00647, 0.00794) | (0.00679, 0.00825) | (0.00695, 0.00845) |
| Mean Dependent Variable | 0.011 | 0.011 | 0.011 |
| Radon Mean | 12.84 | 12.84 | 12.84 |
| 10% Effect Relative to Mean | 6.2% | 6.2% | 6.2% |
| Relative Risk | 1.049 | 1.049 | 1.049 |
| ***Panel B: Veterans in EPA Zones 2 and 3*** | | | |
| Log Cumulative Radon | 0.00598 | 0.00622 | 0.00631 |
| | (0.00572, 0.00640) | (0.00595, 0.00664) | (0.00604, 0.00672) |
| Mean Dependent Variable | 0.0093 | 0.0093 | 0.0093 |
| Radon Mean | 7.72 | 7.72 | 7.72 |
| 10% Effect Relative to Mean | 6.2% | 6.2% | 6.2% |
| Relative Risk | 1.079 | 1.079 | 1.079 |
| **Fixed Effects** | | | |
| Veteran | Yes | Yes | Yes |
| County by Year | Yes | No | No |
| County by 2-Year | No | Yes | No |
| County by 3-Year | No | No | Yes |

Notes: Table reports estimates of 3 using the fixed-effect (within) estimator. Panel A subsets the data to only those veterans living in EPA Zone 1. Panel B subsets the data to only those veterans living in EPA Zones 2 or 3. Each column includes a different set of County and year fixed effects. All regressions include the veteran covariates of smoking status, race, Medicaid eligibility, flag for receiving VA disability and/or receiving social security benefits, agent orange exposure, combat service, sex, second-order polynomials of BMI and age, yearly weather (precipitation, min and max temperatures, min and max dew point temperatures, vapor pressure deficit). Bootstrapped 95 precent confidence intervals are given in parenthesis.

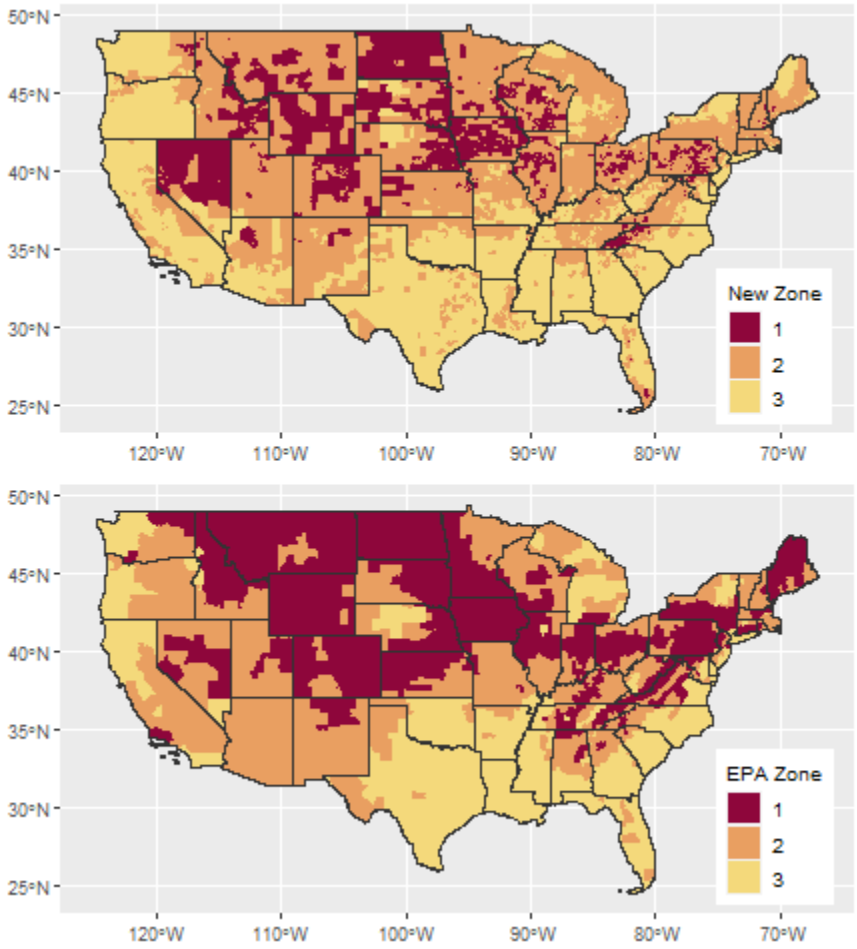
### 0.7.2 Supplementary Figures

Figure 2: Raw Atmospheric Bi-214 Compared to Their Predicted Values




This figure presents the orignal observed values of atmospheric bi-214 in the NURE data as a reference in the top map. The bottom map presents the predicted values after applying the semi-parametric modeling approach in Watkins (2024) The final model is evaluated at the long-term weather normals so that the point predictions can be compared based on the mean distribution of weather across the US.

Figure 3: A New Radon Map at the Census Tract and the EPA Map of Radon Zones




Notes: This figure provides a comparison between the new predictive map of long-term indoor radon from Watkins (2024) and the original EPA Map of Radon Zones. In the top figure, the "New Zone" is assigned based on the average long-term prediction at the census tract level. Census tracts with average long-term predictions exceeding 4 piC/L are classified as New Zone 1, those with average long-term predictions ranging from 2 to 4 piC/L are designated as New Zone 2, and census tracts with average long-term predictions below 2 piC/L are labeled as New Zone 3. The EPA Zones 1-3 are defined similarly, except they represent "probable" screening levels of indoor radon for an area.

## References


Blackmore, D., Herman, M., and Woodward, J. (1982). Heavy gas dispersion models. *Journal of Hazardous Materials*, 6:107–128.

Cao, X. et al. (2017). Radon-induced lung cancer deaths may be overestimated due to failure to account for confounding exposure to diesel engine exhaust in beir vi miner studies. *PLoS ONE*, 12(9).

Chen, E. S., Egger, S., Hughes, S., Müller, N., Schindler, C., Rochat, T., and et al. (2021). Systematic review and meta-analysis of residential radon and lung cancer in never-smokers. *European Respiratory Review*, 30(162):200230.

Darby, S., Hill, C., and Auvinen, A. (1998). Risk of lung cancer associated with residential radon exposure in south-west england: a case-control study. *British Journal of Cancer*, 78(3):394–408.

Darby, S., Hill, C., Auvinen, A., Barros-Dios, J. M., Baysson, H., Bochicchio, F., and et al. (2005). Radon in homes and risk of lung cancer: Collaborative analysis of individual data from 13 european case-control studies. *BMJ*, 330(7485):223–226.

Darby, S., Hill, C., Auvinen, A., Barros-Dios, J. M., Baysson, H., Bochicchio, F., and et al. (2006). Residential radon and lung cancer: detailed results of a collaborative analysis of individual data on 7148 subjects with lung cancer and 14208 subjects without lung cancer from 13 epidemiologic studies in europe. *Scandinavian Journal of Work, Environment & Health*, 32(Suppl 1):1–83.

Field, R. W. et al. (2000). Indoor radon and lung cancer: re-analyses of pooled data from studies in europe. *Health Physics*, 79(6):636–642.

Hu, J., Liu, Y., and Yu, Y. (1994). Estimation of latency period of lung cancer. *Zhonghua Zhong Liu Za Zhi*, 16(1):18–21.

Kreuzer, M., Heinrich, J., Kreienbrock, L., Rosario, A. S., Gerken, M., and Wichmann, H.-E. (2002). Risk factors for lung cancer among nonsmoking women. *International Journal of Cancer*, 100(6):706–713.

Krewski, D., Lubin, J. H., Zielinski, J. M., Alavanja, M., Catalan, V. S., Field, R. W., and et al. (2005). Residential radon and risk of lung cancer: A combined analysis of 7 north american case-control studies. *Epidemiology*, 16(2):137–145.

Krewski, D., Lubin, J. H., Zielinski, J. M., Alavanja, M., Catalan, V. S., Field, R. W., and et al. (2006). A combined analysis of north american case-control studies of residential radon and lung cancer. *Journal of Toxicology and Environmental Health Part A*, 69(7):533–597.

Lehnert, A. and Kearfott, K. (2010). An equilibrium-based model for measuring environmental radon using charcoal canisters. *Health Phys*, 99(2):S154–S163.

Li, C., Wang, C., Yu, J., Fan, Y., Liu, D., Zhou, W., and Shi, T. (2020). Residential radon and histological types of lung cancer: A meta-analysis of case-control studies. *International Journal of Environmental Research and Public Health*, 17(5):1457.

Lipfert, F. W. and Wyzga, R. E. (2019). Longitudinal relationships between lung cancer mortality rates, smoking, and ambient air quality: a comprehensive review and analysis. *Critical Reviews in Toxicology*, 49(9):790–818.

Lubin, J. H. et al. (1994). Radon and lung cancer risk: A joint analysis of 11 underground miner studies. *Journal of the National Cancer Institute*, 86(22):1739–1750.

Malinovsky, G., Yarmoshenko, I., and Vasilyev, A. (2018). Meta-analysis of case-control studies on the relationship between lung cancer and indoor radon exposure. *Radiation and Environmental Biophysics*, 58(1):39–47.

Nazaroff, W. W. (1992). Radon transport from soil to air. *Review of Geophysics*, 30(2):137–160.

Pavia, M., Bianco, A., Pileggi, C., and Angelillo, I. F. (2003). Meta-analysis of residential exposure to radon gas and lung cancer. *Bulletin of the World Health Organization*, 81(10):732–738.

Sandler, D., Weinberg, C., Shore, D., Archer, V., Stone, M., Lyon, J., Rothney-Kozlak, L., Shepherd, M., and Stolwijk, J. (2006). Indoor radon and lung cancer risk in connecticut and utah. *Journal of Toxicology and Environmental Health Part A*, 69(7):633–654.

Scott, B. R. (2011). Residential radon appears to prevent lung cancer. *Dose Response*, 9(4):444–464. Epub 2011 Oct 14.

Watkins, L. (2024). Revaluating the us population's exposure to indoor radon: A model of household indoor radon among 74 million single-family homes.

Wilcox, H., Al-Zoughool, M., Garner, M., Jiang, H., Klotz, J., Krewski, D., Nicholson, W., Schoenberg, J., Villeneuve, P., and Zielinski, J. (2008). Case-control study of radon and lung cancer in new jersey. *Radiation Protection Dosimetry*, 128(2):169–179.