

Inverse probability weighting to estimate impacts of hypothetical occupational limits on radon exposure to reduce lung cancer

Alexander P. Keil, Yi Li, Qing Lan, Stephen Bertke, Robert D. Daniels, Jessie K. Edwards, Kaitlin Kelly-Reif

Contents

APPENDIX S1: Intuition for how the clone-censor-weighting approach addresses HWSB.....	3
APPENDIX S2: Artificial censoring with inverse probability of censoring weighting contrasted with inverse probability of exposure weighting	5
APPENDIX S3: Dynamic treatment regimes considered in the study	6
APPENDIX S4: Illustrative computer code for clone-censor-weight approach	7

List of figures

FIGURE S1: DIRECTED ACYCLIC GRAPHS REPRESENTING A) A SIMPLIFIED VERSION OF THE OBSERVED DATA SUBJECT WITH CONFOUNDING OF THE MONTHLY EXPOSURES (A_1 , A_2 = EXPOSURE AT MONTHS 1 AND 2) EFFECT ON SUBSEQUENT MORTALITY D_2 BY EMPLOYMENT STATUS (E_2) THROUGH UNMEASURED HEALTH STATUS (U) B) OBSERVED DATA AFTER INTRODUCING A TIME-VARYING INDICATOR OF ARTIFICIAL CENSORING (C_1 , C_2) DUE TO EXCEEDING A PERSONAL EXPOSURE LIMIT AND C) WEIGHTED DATA AFTER APPLYING INVERSE PROBABILITY WEIGHTS AND EXCLUDING ALL ARTIFICIALLY CENSORED OBSERVATIONS, WHICH REMOVES CONFOUNDING BY EMPLOYMENT STATUS. OTHER TIME-VARYING AND TIME-FIXED CONFOUNDERS ARE OMITTED FOR CLARITY.	3
FIGURE S2: DISTRIBUTION OF BOOTSTRAP SAMPLE ESTIMATES (HISTOGRAMS) CONTRASTED WITH NORMAL DISTRIBUTIONS FOR A REPRESENTATIVE SET OF EFFECT ESTIMATES. THE TOP ROW CONTAINS LN(HAZARD RATIO) ESTIMATES WHILE THE BOTTOM CONTAINS RISK DIFFERENCE ESTIMATES ACROSS 200 BOOTSTRAP SAMPLES. THE LEFT TWO COLUMNS CONTAIN ESTIMATES FROM THE NON-POOLED APPROACHES, AND THE RIGHT TWO COLUMNS CONTAIN ESTIMATES FROM THE MARGINAL STRUCTURAL MODEL. THE FIRST AND THIRD COLUMNS REPRESENT EFFECT MEASURES FOR THE 10 WLM LIMIT, AND THE SECOND AND FOURTH COLUMNS REPRESENT EFFECT MEASURES FOR THE 1 WLM LIMIT.	9
FIGURE S3: CUMULATIVE LUNG CANCER MORTALITY RISK ESTIMATES AND 95% POINTWISE BOOTSTRAP CONFIDENCE INTERVALS AT 90 AMONG 4,127 WHITE AND NATIVE AMERICAN MINERS FROM THE COLORADO PLATEAU URANIUM MINERS COHORT. POINTS AND WHISKERS REPRESENT NON-PARAMETRIC WEIGHTED ESTIMATES AND 95% CONFIDENCE INTERVALS, WHILE THE DASHED LINE AND GRAY BAND REPRESENT THE POINT ESTIMATE AND 95% CONFIDENCE INTERVALS DERIVED FROM ESTIMATES OF MARGINAL STRUCTURAL COX MODEL PARAMETERS COMBINING ALL DYNAMIC REGIMES, WHICH WAS IN TURN USED TO ESTIMATE THE BASELINE HAZARD AT ZERO RADON EXPOSURE AND CUMULATIVE RISK.	10
FIGURE S4: WEIGHTED, NON-PARAMETRIC ESTIMATES OF THE RISK DIFFERENCE CONTRASTING TWO DYNAMIC REGIMES WITH RISK UNDER NO INTERVENTION AMONG 4,127 WHITE AND NATIVE AMERICAN MINERS FROM THE COLORADO PLATEAU URANIUM MINERS COHORT.	11

List of tables

TABLE S1: CHARACTERISTICS OF OBSERVED DATA AND DATA UNDER A SET OF 8 REPRESENTATIVE COPIES (OUT OF 34) OF THE OBSERVED DATA WITH ARTIFICIAL CENSORING, ALONG WITH INVERSE PROBABILITY OF CENSORING WEIGHT DISTRIBUTIONS ACROSS PERSON TIME IN EACH COPY OF THE DATA, 4,127 WHITE AND NATIVE AMERICAN MINERS FROM THE COLORADO PLATEAU URANIUM MINERS COHORT. SENSITIVITY ANALYSIS: ADD AGE-AT-HIRE TO CENSORING MODELS (COMPARE WITH TABLE 2 IN THE MAIN TEXT).	12
---	----

TABLE S2: SENSITIVITY ANALYSIS: CAUSE SPECIFIC RISK OF LUNG CANCER THROUGH AGE 90, RISK DIFFERENCES AND HAZARD RATIOS CONTRASTING MULTIPLE DYNAMIC REGIMES USING NON-PARAMETRIC OR SEMI-PARAMETRIC ESTIMATORS FOR 8 REPRESENTATIVE DYNAMIC REGIMES AMONG 4,127 WHITE AND NATIVE AMERICAN MINERS FROM THE COLORADO PLATEAU URANIUM MINERS COHORT. SENSITIVITY ANALYSIS: ADD AGE-AT-HIRE TO CENSORING MODELS (COMPARE WITH TABLE 3 IN THE MAIN TEXT).	13
TABLE S3: CHARACTERISTICS OF OBSERVED DATA AND DATA UNDER A SET OF 8 REPRESENTATIVE COPIES (OUT OF 34) OF THE OBSERVED DATA WITH ARTIFICIAL CENSORING, ALONG WITH INVERSE PROBABILITY OF CENSORING WEIGHT DISTRIBUTIONS ACROSS PERSON TIME IN EACH COPY OF THE DATA, 4,127 WHITE AND NATIVE AMERICAN MINERS FROM THE COLORADO PLATEAU URANIUM MINERS COHORT. SENSITIVITY ANALYSIS: ADD AGE-AT-HIRE AND CUMULATIVE YEARS OF EMPLOYMENT TO CENSORING MODELS (COMPARE WITH TABLE 2 IN THE MAIN TEXT).	14
TABLE S4: CAUSE SPECIFIC RISK OF LUNG CANCER THROUGH AGE 90, RISK DIFFERENCES AND AND HAZARD RATIOS CONTRASTING MULTIPLE DYNAMIC REGIMES USING NON-PARAMETRIC OR SEMI-PARAMETRIC ESTIMATORS FOR 8 REPRESENTATIVE DYNAMIC REGIMES AMONG 4,127 WHITE AND NATIVE AMERICAN MINERS FROM THE COLORADO PLATEAU URANIUM MINERS COHORT. SENSITIVITY ANALYSIS: ADD AGE-AT-HIRE AND CUMULATIVE YEARS OF EMPLOYMENT TO CENSORING MODELS (COMPARE WITH TABLE 3 IN THE MAIN TEXT).	15
TABLE S5: CHARACTERISTICS OF OBSERVED DATA AND DATA UNDER A SET OF 8 REPRESENTATIVE COPIES (OUT OF 34) OF THE OBSERVED DATA WITH ARTIFICIAL CENSORING, ALONG WITH INVERSE PROBABILITY OF CENSORING WEIGHT DISTRIBUTIONS ACROSS PERSON TIME IN EACH COPY OF THE DATA, 4,127 WHITE AND NATIVE AMERICAN MINERS FROM THE COLORADO PLATEAU URANIUM MINERS COHORT. SENSITIVITY ANALYSIS ADD AGE-AT-HIRE AND ANNUAL EXPOSURE (2 YEAR LAG) TO CENSORING MODELS (COMPARE WITH TABLE 2 IN THE MAIN TEXT).	16
TABLE S6: CAUSE SPECIFIC RISK OF LUNG CANCER THROUGH AGE 90, RISK DIFFERENCES AND AND HAZARD RATIOS CONTRASTING MULTIPLE DYNAMIC REGIMES USING NON-PARAMETRIC OR SEMI-PARAMETRIC ESTIMATORS FOR 8 REPRESENTATIVE DYNAMIC REGIMES AMONG 4,127 WHITE AND NATIVE AMERICAN MINERS FROM THE COLORADO PLATEAU URANIUM MINERS COHORT. SENSITIVITY ANALYSIS ADD AGE-AT-HIRE AND ANNUAL EXPOSURE (2 YEAR LAG) TO CENSORING MODELS (COMPARE WITH TABLE 3 IN THE MAIN TEXT).	17

APPENDIX S1: Intuition for how the clone-censor-weighting approach addresses HWSB

The clone-censor-weight approach can be summarized as a series of directed acyclic graphs (**Figure S1**). First, exposure-mortality effect estimates in the observed data are subject to confounding by employment status, but adjustment for employment status through regression would introduce collider stratification bias (i.e., HWSB, panel A). Next, exposure levels are used to deterministically (denoted by a double box around a node) censor individuals (Panel B). After weighting based on a model for censoring, past levels of confounders no longer predict censoring, which effectively removes confounding by time-fixed and time-varying variables in the weighted population because censoring was a deterministic function of exposure, so that arrows can be removed from exposure (Panel C). Thus, in the final weighted dataset, analyses of the weighted data are not subject to HWSB. However, we also note that HWSB is structurally identical to other settings for which the clone-censor-weight approach was developed, such as questions about the optimal threshold of CD4 count at which to start anti-retroviral treatment in HIV patients (1, 2). In more general terms, a heuristic for inverse probability weighting is that the weighted population represents the target population (here, the observed data), had everyone been exposed to a specific exposure level. The weighted “pseudo-population” shares the same confounder distribution as the target population, so that confounders no longer predict exposure levels when multiple pseudo-populations are compared and thus no longer operate as confounders (e.g. when estimating a risk difference contrasting two exposure levels).

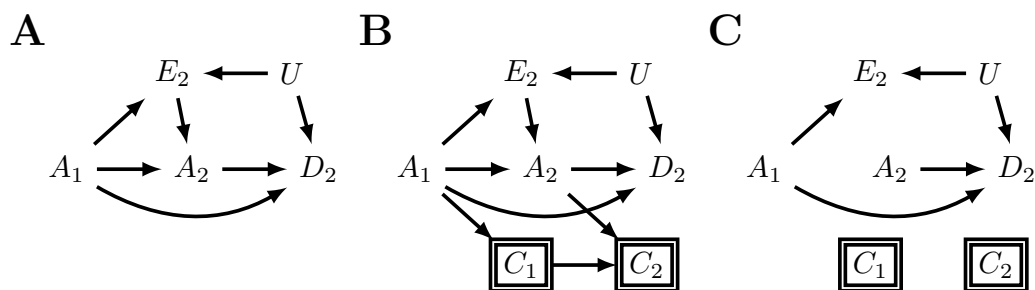


Figure S1: Directed acyclic graphs representing A) a simplified version of the observed data subject with confounding of the monthly exposures (A_1, A_2 = exposure at months 1 and 2) effect on subsequent mortality D_2 by employment status (E_2) through unmeasured health status (U) B) observed data after introducing a time-varying indicator of artificial censoring (C_1, C_2) due to exceeding a personal exposure limit and C) weighted data after applying inverse probability weights and excluding all artificially censored observations, which removes confounding by employment status. Other time-varying and time-fixed confounders are omitted for clarity.

If we ran a trial where exposure to a specific regime was randomized at baseline (e.g. “if at work, have a uniform random exposure below 2 WLM”), then we could calculate the causal risk difference between the two arms using an unadjusted approach, like Kaplan-Meier. However, if healthier trial members could have higher cumulative exposures, then we could not get a causal estimate of the cumulative-exposure-response curve because cumulative exposure would be subject to confounding in a way that the baseline exposure regimes are not. The crude risk difference would be equivalent to an “intent-to-treat” analysis. In that trial, we would not even need to consider a censoring model for $m > 1$, and there would be no bias for the effect of randomization to a specific arm (i.e. the effect of having a specific occupational limit), even though HWSB would still be present if we attempted to estimate the effect of cumulative exposure. If assignment to a given arm were conditionally assigned (e.g. smoking and non-smoking workers were assigned with different probabilities to different arms), then we might wish to apply inverse probability of censoring weights to address the imbalance in smoking, since that imbalance could otherwise explain differences across trial arms in the absence of an effect of exposure. The clone-censor-weight approach

weights data in such a way that we have balance over covariates and can analyze the weighted data like this simple randomized trial

Now, consider that same trial with the additional wrinkle that some individuals will randomly have exposures above the assigned limit and thus are no longer compliant with the study. Now, because some workers are non-compliant, the crude analysis would be biased because (if exposure is harmful), mortality would be higher than expected due to non-compliance. If non-compliance is perfectly random (e.g. every person-year has the same probability of non-compliance), we would be free to censor those individuals when the exposure exceeds the assigned limit and conduct a crude Kaplan-Meier analysis on the censored data, where that analysis encodes the assumption that the censored trial patients have the same hazard as the uncensored trial patients. If non-compliance is not perfectly random (say, smokers are more likely to have exposures that exceed the limit), then we would need to create inverse-probability of censoring weights that depended on smoking status, which would remove the imbalance of smoking vs. non-smoking in terms of remaining in the study at each censoring time. Note that healthy worker survivor bias is not directly addressed in this case (the $m > 1$ censoring case), because the phenomenon that requires the censoring model has to do with the conditional probability of exposure exceeding an exposure limit. The probability of exposure, conditional on baseline and time-dependent covariates, is certainly involved in HWSB as described in Figure 2A. If exposure does not depend on employment, then HWSB is not an issue since cumulative exposure would no longer be determined by duration of employment.

Finally, consider the same trial but with a different estimand like “have a uniform random exposure greater than 0 and less than 2 WLM both on and off work.” In that case, employment becomes relevant in the trial because those who leave employment would be unexposed and out of compliance. We could then use IPCW with weights based on the probability of leaving employment (and therefore artificial censoring due to non-compliance), which would allow individuals remaining on the study (employed workers) to represent individuals who left work, had they, counter to fact, remained exposed. However, in that case, we would need to identify the common causes of leaving employment and mortality. This would be equivalent to measuring U in diagram Figure 2A, which we assume we do not have sufficient data to address since we consider U to be multi-factorial and strongly represented by latent underlying health processes with questionable measureability. For the same estimand, we could not use inverse probability of exposure weighting because of the problem of non-positivity, where we would want to include employment in the exposure model but then could not estimate a positive probability of exposure among the unemployed.

In summary, HWSB is in some regard addressed by the study question itself and the use of inverse probability weighting simply allows us to get an internally valid answer to that study question (under the identification assumptions).

REFERENCES

1. Cain LE, Saag MS, Petersen M, et al. Using observational data to emulate a randomized trial of dynamic treatment-switching strategies: an application to antiretroviral therapy. *Int J Epidemiol* 2016;45(6):2038-49.
2. Hernan MA, Lanoy E, Costagliola D, Robins JM. Comparison of dynamic treatment regimes via inverse probability weighting. *Basic Clin Pharmacol Toxicol* 2006;98(3):237-42.

APPENDIX S2: Artificial censoring with inverse probability of censoring weighting contrasted with inverse probability of exposure weighting

For a static (i.e. non-dynamic) regimes, inverse probability of treatment weights are often used to contrast different exposure levels. As an example, in a cohort without loss-to-follow-up, for a time-fixed regime “always expose”, the month-specific contribution to the weight for a binary exposure A_m with time-varying confounders E_m for someone who is, in fact always exposed, is given as

$$\frac{1}{\Pr(A_m = 1 \mid \bar{E}_m, \bar{A}_{m-1}, \mathbf{Z})}$$

Estimation of the risk under this regime could be conducted by finding the group of individuals who are always exposed, and estimating a weighted survival curve for this group. This process can be generalized to a setting of a marginal structural model, where we apply weights to all individuals in the population with the time-specific contributions of the weights based on the probability that an individual had their observed exposure value.

Similarly, we can view the inverse probability of not being artificially censored as the equivalent to the inverse probability of following a specific dynamic treatment regime, given by $g_A(\bar{E}_m)$, which can be expressed as a rule such as “if at work, monthly exposures must be under 1 WLM” (or more precisely “remain exposed at some level of exposure that follows the empirical distribution of monthly exposures among the person-time with exposure less than 1 WLM”). Thus, the time-specific inverse probability of censoring weight contribution can be expressed as

$$\frac{1}{1 - \Pr(C_m = 1 \mid \bar{E}_m, \bar{A}_{m-1}, \mathbf{Z})} \equiv \frac{1}{\Pr(A_m \in g_A(\bar{E}_m) \mid \bar{E}_m, \bar{A}_{m-1}, \mathbf{Z})}$$

Thus, artificially censoring of an individual’s person-time after they exceed a specific occupational limit, followed by weighting their remaining person-time by the inverse probability of not being artificially censored is equivalent to weighting by the inverse probability of following a dynamic exposure regime of interest. This equivalence follows from the fact that artificial censoring is a deterministic function of exposure.

Often inverse probability weighting schemes do not employ artificial censoring and instead weight the observed data by the inverse probability that each observation has had the exposure that was observed. A natural question is why that same scheme is not implemented here. The basis for using a different analytic scheme is that a given person-month in the study may be consistent with multiple different dynamic regimes. This results in a given worker effectively having multiple different exposures at the same time for a given analysis.

For example, say we are contrasting mortality under two different occupational limits: 10 WLM and 2 WLM. An observation of a worker whose exposure in their first month is 1.2 WLM is consistent with occupational limits of 2 WLM or 10 WLM (among many others). This particular person month would receive a weight for each of these limits, and those weights would correspond to different dynamic regimes. Here, the “dynamic regime” is the exposure of interest. There is no straightforward way to allow a single observation to have two different exposures in the same dataset, so instead we employ copies of the dataset with artificial censoring as a convenient way to make contrasts of mortality across different dynamic regimes.

APPENDIX S3: Dynamic treatment regimes considered in the study

In units of working level months, the following 34 values represent the dynamic treatment regimes (expressed as limits on cumulative exposure within a month) considered in this study:

0.33, 0.75, 0.83, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, 10.0, 12.0, 14.0, 16.0, 18.0, 20.0, 25.0, 30.0, 35.0, 40.0, 45.0, 50.0, 75.0

Additionally, we considered a 35th regime of “natural course”, which is the dynamic treatment regime that is represented by the empirical distribution of monthly exposures. In these analysis, the CPUM under the “natural course” intervention is simply the observed data.

APPENDIX S4: Illustrative computer code for clone-censor-weight approach

```
#####
## Language: Julia v 1.10 (similar to R syntax)
#####
#### The data are stored in a DataFrame called "data"
# Structure: 1 record per person-year, sorted by age (ascending)
# Variables list
# id: unique identifier
# firstid: 1 = first record per individual, 0 = other
# inage: age at the beginning of the person-year
# outage: age at the end of the person-year
# outage_sp1, outage_sp2: age at the end of the person-year, splines
# d_lc: indicator of lung cancer death in person-year (1=yes,0=no)
# dob: date of birth
# dob_sp1, dob_sp2: date of birth splines
# nonwhite: race (0=white,1=other, primarily Native American)
# smoke: smoking (0=never, 1=other)
# atwork: employment status (1= work during person year, 0= no work during person year)
# wlm: annual exposure in working level months
# cumwlm_lag5: cumulative exposure in working level months
# cumwlm_lag5_sp1, cumwlm_lag5_sp2: cumulative exposure in working level months, splines
#####

using LSurvival # available at https://github.com/alexpkeil1/LSurvival.jl
using GLM, DataFrames
# decide on annual limit
LIMIT = 10 # 10 working level months in a year

# clone the data
intdata = deepcopy(data)

# create censoring indicators, initialized at zero
intdata.cens_limit .= 0.0
intdata.cens_prev .= 0.0 # observations after censoring: flag for removing from data
for obs in 1:nrow(data)
    if intdata.firstid[obs] == 1
        intdata.cens_prev[obs] = 0
    end
    intdata.cens_limit[obs] = ((intdata.wlm[obs] > LIMIT) & (intdata.cens_prev[obs] == 0)) ? 1 :
0
    if intdata.firstid[obs] == 0
        intdata.cens_prev[obs] = ((intdata.cens_prev[obs-1] == 1) || (intdata.cens_limit[obs-1] ==
1)) ? 1 : 0
    end
end

# restrict for modeling to uncensored observations and first censored observation
intdata = intdata[findall(intdata.cens_prev .== 0),:]

# create a copy of the data at baseline
baseline_data = intdata[findall(intdata.firstid .== 1),:]

# model censoring at baseline (everyone is at work at baseline)
basemod =
glm(@formula(cens_limit~outage+outage_sp1+outage_sp2+dob+dob_sp1+dob_sp2+nonwhite+smoke+cumwlm_la
g5+cumwlm_lag5_sp1+cumwlm_lag5_sp2), baseline_data, Binomial(), LogitLink())
basemodn = glm(@formula(cens_limit~outage+outage_sp1+outage_sp2), baseline_data, Binomial(),
LogitLink())

# model censoring over the rest of the employed follow-up time
workdata = intdata[findall((intdata.atwork .== 1) .& (intdata.firstid .== 0)),:]
fumod =
glm(@formula(cens_limit~outage+outage_sp1+outage_sp2+dob+dob_sp1+dob_sp2+nonwhite+smoke+cumwlm_la
g5+cumwlm_lag5_sp1+cumwlm_lag5_sp2), workdata, Binomial(), LogitLink())
```

```

fumodn = glm(@formula(cens_limit~outage+outage_sp1+outage_sp2), workdata, Binomial(),
LogitLink())

# initialize weight numerator and denominator at 1.0
intdata.num .= 1.0
intdata.den .= 1.0
intdata.wt .= 1.0
# now calculate the weights over time,
# 1. start with numerator and denominator from the "baseline" model
intdata.num[findall(intdata.firstid .== 1)] = 1.0 .- predict(basemodn)
intdata.den[findall(intdata.firstid .== 1)] = 1.0 .- predict(basemodn)
# 2. numerator and denominator over follow-up
intdata.num[findall((intdata.firstid .== 0) .& (intdata.atwork .== 1))] = 1.0 .- predict(fumodn)
intdata.den[findall((intdata.firstid .== 0) .& (intdata.atwork .== 1))] = 1.0 .- predict(fumodn)

# calculate weights
for obs in 1:nrow(intdata)
    if intdata.firstid[obs] == 1
        intdata.wt[obs] = 1.0
    end
    if intdata.firstid[obs] == 0
        intdata.wt[obs] = intdata.wt[obs-1] * intdata.num[obs] / intdata.den[obs]
    end
end

# now drop the remaining censored observations, which are no longer needed for modeling
intdata = intdata[findall(intdata.cens_limit .== 0),:]

# now we have a single "pseudo-dataset", we can fit non-parametric survival methods
# (note Aalen-Johansen estimator requires a competing risk – here we use Kaplan-Meier
# for simplicity in illustration)

kaplan_meier(@formula(Surv(inage, outage, d_lc)~1), intdata, wts=intdata.wt)

#####
# estimating a cause-specific hazard ratio contrasting a limit of 10 WLM with the natural course
#####
#combine intervention data with the original data (which had virtually zero loss-to-follow-up and
thus no need to model censoring by loss-to-follow-up)

intdata.limit .= "10"
data.limit .= "None"
data.wt .= 1.0 # original data get a weight of 1.0
joint_data = vcat(data, intdata, cols=:intersect)

# fit a weighted Cox model with indicator term for exposure
coxph(@formula(Surv(inage, outage, d_lc)~limit), joint_data, ties="efron", contrasts =
Dict{:limit => DummyCoding(base="None")}, wts=joint_data.wt)

# this model can form the basis of risk estimates after estimating the baseline hazard and
accounting for competing risks (not shown here)

```

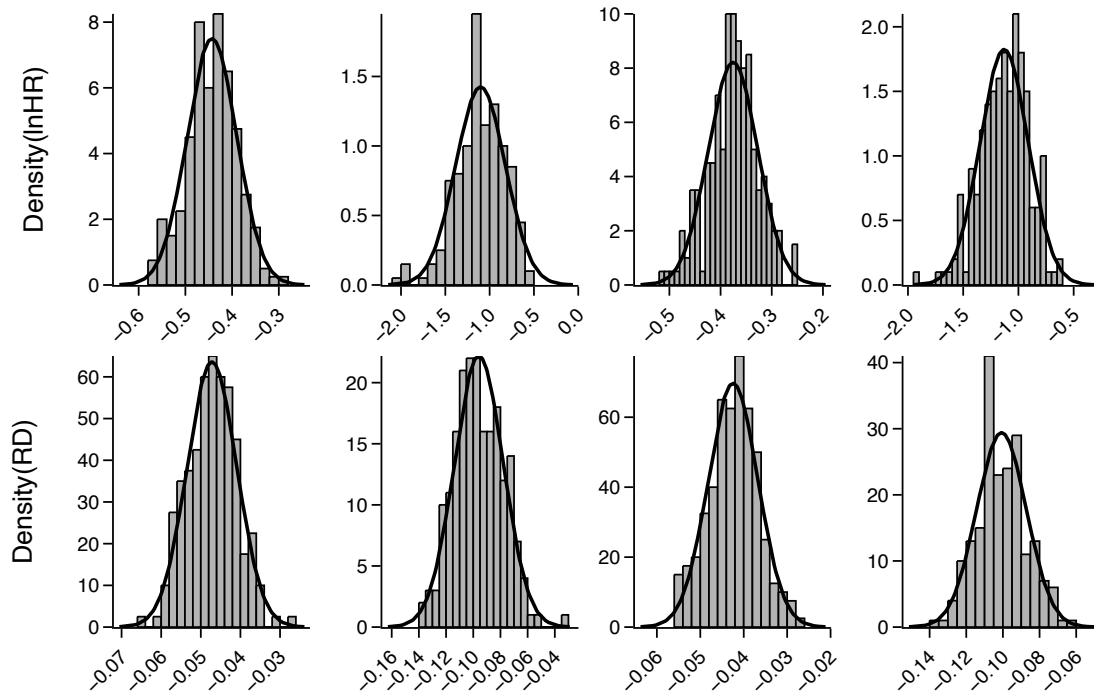


Figure S2: Distribution of bootstrap sample estimates (histograms) contrasted with normal distributions for a representative set of effect estimates. The top row contains $\ln(\text{hazard ratio})$ estimates while the bottom contains risk difference estimates across 200 bootstrap samples. The left two columns contain estimates from the non-pooled approaches, and the right two columns contain estimates from the

marginal structural model. The first and third columns represent effect measures for the 10 WLM limit, and the second and fourth columns represent effect measures for the 1 WLM limit.

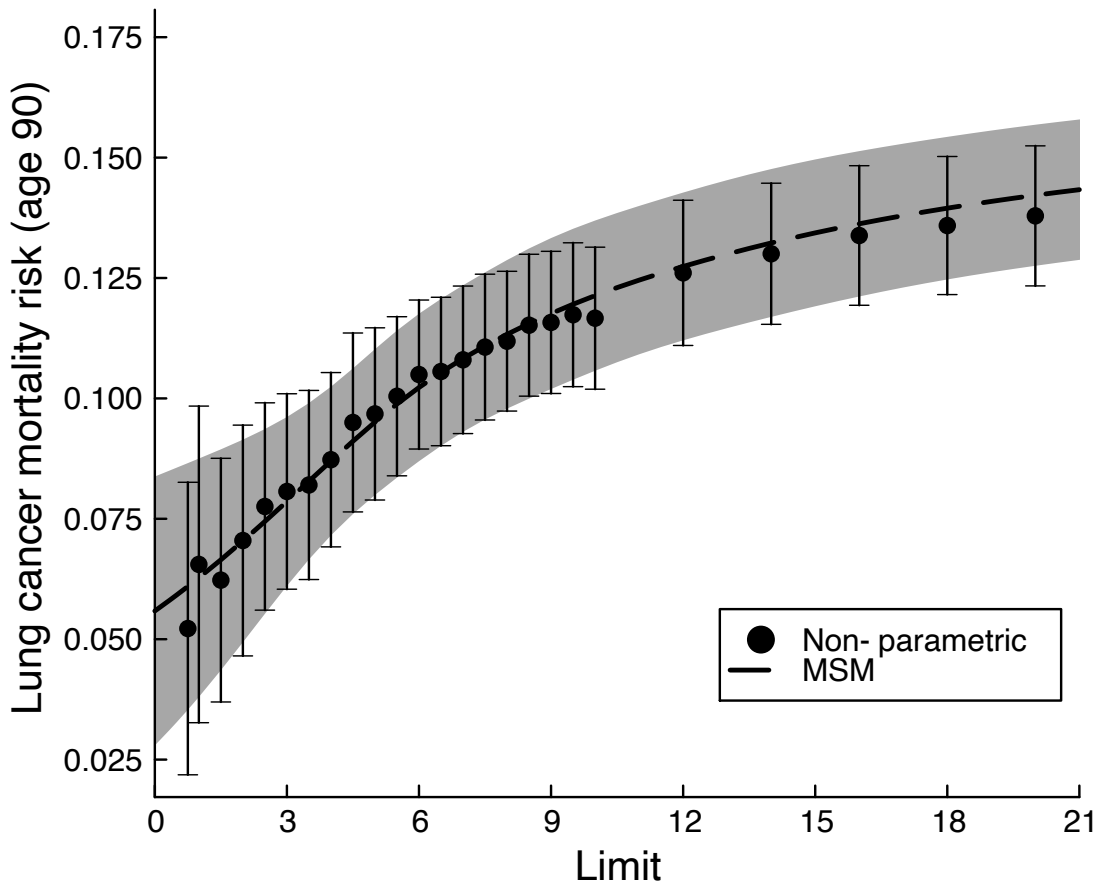


Figure S3: Cumulative lung cancer mortality risk estimates and 95% pointwise bootstrap confidence intervals at 90 among 4,127 White and Native American miners from the Colorado Plateau Uranium Miners cohort. Points and whiskers represent non-parametric weighted estimates and 95% confidence intervals, while the dashed line and gray band represent the point estimate and 95% confidence intervals derived from estimates of marginal structural Cox model parameters combining all dynamic regimes, which was in turn used to estimate the baseline hazard at zero radon exposure and cumulative risk.

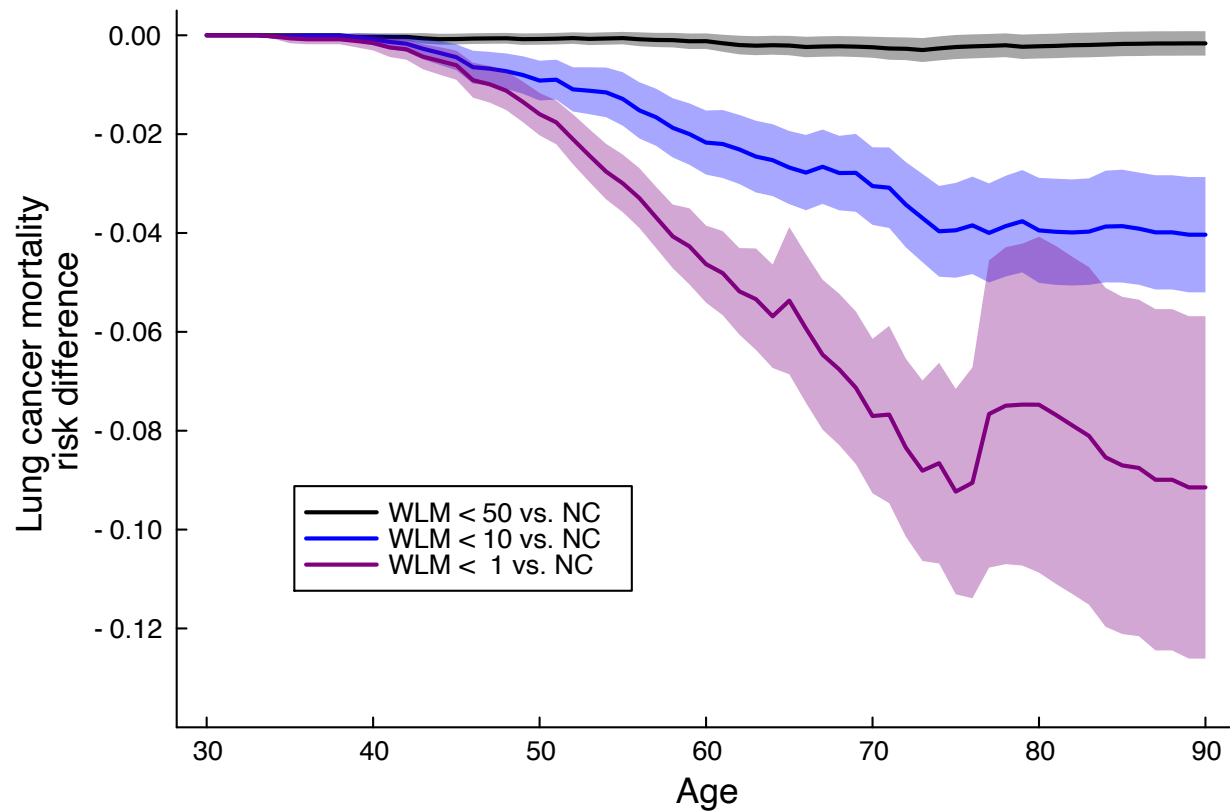


Figure S4: Weighted, non-parametric estimates of the risk difference contrasting two dynamic regimes with risk under no intervention among 4,127 White and Native American miners from the Colorado Plateau Uranium Miners cohort.

Table S1: Characteristics of observed data and data under a set of 8 representative copies (out of 34) of the observed data with artificial censoring, along with inverse probability of censoring weight distributions across person time in each copy of the data, 4,127 White and Native American miners from the Colorado Plateau Uranium Miners cohort. Sensitivity analysis: add age-at-hire to censoring models (compare with Table 2 in the main text).

Limit	Lung cancer deaths	Artificially censored		Person-years of follow-up (×1000)	Inverse probability weights
	N (%)	At baseline N (%)	During follow-up N (%)		Mean (min, max)
None	613 (100%)	0 (0%)	0 (0%)	134.6	1.00 (1.00, 1.00)
75 WLM	566 (92%)	78 (2%)	160 (4%)	127.6	1.00 (0.88, 2.25)
50 WLM	541 (88%)	150 (4%)	220 (5%)	123.8	1.00 (0.84, 2.50)
20 WLM	383 (62%)	705 (17%)	489 (12%)	98.5	1.01 (0.61, 4.27)
10 WLM	249 (41%)	1327 (32%)	617 (15%)	74.6	1.02 (0.46, 9.52)
5 WLM	124 (20%)	2133 (52%)	607 (15%)	48.8	1.01 (0.23, 5.31)
2 WLM	36 (6%)	3117 (76%)	379 (9%)	22.4	0.99 (0.19, 3.94)
1 WLM	17 (3%)	3549 (86%)	193 (5%)	13.2	1.01 (0.10, 4.07)
0.75 WLM	13 (2%)	3641 (88%)	149 (4%)	11.5	1.08 (0.06, 6.27)
0.33 WLM	10 (2%)	3821 (93%)	67 (2%)	7.7	1.85 (0.02, 59.09)

Table S2: Sensitivity analysis: Cause specific risk of lung cancer through age 90, risk differences and hazard ratios contrasting multiple dynamic regimes using non-parametric or semi-parametric estimators for 8 representative dynamic regimes among 4,127 White and Native American miners from the Colorado Plateau Uranium Miners cohort. Sensitivity analysis: add age-at-hire to censoring models (compare with Table 3 in the main text).

Limit	Non-parametric			Marginal Structural model		
	Age-90 risk (95% CI)	Age-90 risk difference (95% CI)	Hazard ratio (95% CI)	Age-90 risk (95% CI)	Age-90 risk difference (95% CI)	Hazard ratio (95% CI)
None	0.16 (0.15, 0.18)	0.01 (0.00, 0.01)				
75 WLM	0.16 (0.14, 0.17)	Ref	Ref	0.16 (0.15, 0.17)	Ref	Ref
50 WLM	0.15 (0.14, 0.17)	0.00 (0.00, 0.00)	0.98 (0.96, 1.00)	0.16 (0.15, 0.17)	0.00 (-0.01, 0.00)	0.98 (0.96, 1.00)
20 WLM	0.14 (0.12, 0.15)	-0.02 (-0.03, -0.01)	0.83 (0.78, 0.88)	0.14 (0.13, 0.15)	-0.02 (-0.03, -0.01)	0.83 (0.79, 0.88)
10 WLM	0.11 (0.10, 0.13)	-0.04 (-0.05, -0.03)	0.67 (0.60, 0.74)	0.12 (0.10, 0.13)	-0.04 (-0.05, -0.03)	0.68 (0.62, 0.75)
5 WLM	0.09 (0.08, 0.11)	-0.06 (-0.08, -0.05)	0.51 (0.44, 0.60)	0.09 (0.08, 0.11)	-0.07 (-0.08, -0.05)	0.52 (0.45, 0.60)
2 WLM	0.07 (0.04, 0.09)	-0.09 (-0.11, -0.07)	0.37 (0.26, 0.51)	0.07 (0.05, 0.09)	-0.10 (-0.12, -0.07)	0.36 (0.27, 0.49)
1 WLM	0.06 (0.03, 0.09)	-0.10 (-0.13, -0.06)	0.34 (0.18, 0.61)	0.06 (0.04, 0.08)	-0.10 (-0.13, -0.08)	0.32 (0.21, 0.48)
0.75 WLM	0.04 (0.02, 0.07)	-0.11 (-0.14, -0.08)	0.27 (0.13, 0.57)	0.06 (0.03, 0.08)	-0.11 (-0.13, -0.08)	0.30 (0.19, 0.47)
0.33 WLM				0.05 (0.03, 0.08)	-0.11 (-0.14, -0.08)	0.29 (0.17, 0.47)
No exposure				0.05 (0.03, 0.08)	-0.11 (-0.14, -0.08)	0.27 (0.16, 0.47)

Table S3: Characteristics of observed data and data under a set of 8 representative copies (out of 34) of the observed data with artificial censoring, along with inverse probability of censoring weight distributions across person time in each copy of the data, 4,127 White and Native American miners from the Colorado Plateau Uranium Miners cohort. Sensitivity analysis: add age-at-hire and cumulative years of employment to censoring models (compare with Table 2 in the main text).

Limit	Lung cancer deaths	Artificially censored		Person-years of follow-up (X1000)	Inverse probability weights
	N (%)	At baseline N (%)	During follow-up N (%)		Mean (min, max)
None	613 (100%)	0 (0%)	0 (0%)	134.6	1.00 (1.00, 1.00)
75 WLM	566 (92%)	78 (2%)	160 (4%)	127.6	1.00 (0.88, 2.18)
50 WLM	541 (88%)	150 (4%)	220 (5%)	123.8	1.00 (0.83, 2.60)
20 WLM	383 (62%)	705 (17%)	489 (12%)	98.5	1.01 (0.59, 4.62)
10 WLM	249 (41%)	1327 (32%)	617 (15%)	74.6	1.02 (0.45, 9.61)
5 WLM	124 (20%)	2133 (52%)	607 (15%)	48.8	1.01 (0.29, 6.71)
2 WLM	36 (6%)	3117 (76%)	379 (9%)	22.4	1.01 (0.19, 7.03)
1 WLM	17 (3%)	3549 (86%)	193 (5%)	13.2	0.98 (0.08, 6.12)
0.75 WLM	13 (2%)	3641 (88%)	149 (4%)	11.5	0.97 (0.05, 7.18)
0.33 WLM	10 (2%)	3821 (93%)	67 (2%)	7.7	1.01 (0.02, 21.83)

Table S4: Cause specific risk of lung cancer through age 90, risk differences and hazard ratios contrasting multiple dynamic regimes using non-parametric or semi-parametric estimators for 8 representative dynamic regimes among 4,127 White and Native American miners from the Colorado Plateau Uranium Miners cohort. Sensitivity analysis: add age-at-hire and cumulative years of employment to censoring models (compare with Table 3 in the main text).

Limit	Non-parametric			Marginal Structural model		
	Age-90 risk (95% CI)	Age-90 risk difference (95% CI)	Hazard ratio (95% CI)	Age-90 risk (95% CI)	Age-90 risk difference (95% CI)	Hazard ratio (95% CI)
None	0.16 (0.15, 0.18)	0.01 (0.00, 0.01)				
75 WLM	0.16 (0.14, 0.17)	Ref	Ref	0.16 (0.15, 0.17)	Ref	Ref
50 WLM	0.15 (0.14, 0.17)	0.00 (0.00, 0.00)	0.98 (0.96, 1.00)	0.16 (0.15, 0.17)	0.00 (-0.01, 0.00)	0.97 (0.96, 0.99)
20 WLM	0.13 (0.12, 0.15)	-0.02 (-0.03, -0.01)	0.83 (0.78, 0.88)	0.14 (0.13, 0.15)	-0.02 (-0.03, -0.02)	0.82 (0.78, 0.87)
10 WLM	0.12 (0.10, 0.13)	-0.04 (-0.05, -0.03)	0.67 (0.60, 0.74)	0.12 (0.11, 0.13)	-0.04 (-0.05, -0.03)	0.69 (0.62, 0.76)
5 WLM	0.10 (0.08, 0.11)	-0.06 (-0.08, -0.04)	0.51 (0.44, 0.60)	0.10 (0.08, 0.11)	-0.07 (-0.08, -0.05)	0.54 (0.47, 0.62)
2 WLM	0.07 (0.04, 0.09)	-0.09 (-0.11, -0.06)	0.37 (0.26, 0.51)	0.07 (0.04, 0.09)	-0.10 (-0.12, -0.07)	0.36 (0.26, 0.50)
1 WLM	0.05 (0.02, 0.08)	-0.11 (-0.14, -0.08)	0.34 (0.18, 0.61)	0.06 (0.03, 0.08)	-0.11 (-0.13, -0.08)	0.31 (0.19, 0.48)
0.75 WLM	0.04 (0.01, 0.08)	-0.11 (-0.15, -0.07)	0.27 (0.13, 0.57)	0.05 (0.03, 0.08)	-0.11 (-0.14, -0.08)	0.29 (0.18, 0.47)
0.33 WLM				0.05 (0.02, 0.08)	-0.11 (-0.14, -0.08)	0.27 (0.16, 0.47)
No exposure				0.05 (0.02, 0.08)	-0.11 (-0.14, -0.08)	0.26 (0.15, 0.46)

Table S5: Characteristics of observed data and data under a set of 8 representative copies (out of 34) of the observed data with artificial censoring, along with inverse probability of censoring weight distributions across person time in each copy of the data, 4,127 White and Native American miners from the Colorado Plateau Uranium Miners cohort. Sensitivity analysis add age-at-hire and annual exposure (2 year lag) to censoring models (compare with Table 2 in the main text).

Limit	Lung cancer deaths	Artificially censored		Person-years of follow-up (X1000)	Inverse probability weights
	N (%)	At baseline N (%)	During follow-up N (%)		Mean (min, max)
None	613 (100%)	0 (0%)	0 (0%)	134.6	1.00 (1.00, 1.00)
75 WLM	566 (92%)	78 (2%)	160 (4%)	127.6	1.00 (0.88, 2.34)
50 WLM	541 (88%)	150 (4%)	220 (5%)	123.8	1.00 (0.84, 2.55)
20 WLM	383 (62%)	705 (17%)	489 (12%)	98.5	1.01 (0.61, 4.29)
10 WLM	249 (41%)	1327 (32%)	617 (15%)	74.6	1.03 (0.44, 9.32)
5 WLM	124 (20%)	2133 (52%)	607 (15%)	48.8	1.02 (0.20, 9.44)
2 WLM	36 (6%)	3117 (76%)	379 (9%)	22.4	1.01 (0.12, 13.22)
1 WLM	17 (3%)	3549 (86%)	193 (5%)	13.2	1.22 (0.05, 13.80)
0.75 WLM	13 (2%)	3641 (88%)	149 (4%)	11.5	1.45 (0.04, 35.91)
0.33 WLM	10 (2%)	3821 (93%)	67 (2%)	7.7	1.98 (0.02, 85.52)

Table S6: Cause specific risk of lung cancer through age 90, risk differences and hazard ratios contrasting multiple dynamic regimes using non-parametric or semi-parametric estimators for 8 representative dynamic regimes among 4,127 White and Native American miners from the Colorado Plateau Uranium Miners cohort. Sensitivity analysis add age-at-hire and annual exposure (2 year lag) to censoring models (compare with Table 3 in the main text).

Limit	Non-parametric			Marginal Structural model		
	Age-90 risk (95% CI) ^a	Age-90 risk difference (95% CI)	Hazard ratio (95% CI)	Age-90 risk (95% CI)	Age-90 risk difference (95% CI)	Hazard ratio (95% CI)
None	0.16 (0.15, 0.18)	0.01 (0.00, 0.01)				
75 WLM	0.16 (0.14, 0.17)	Ref	Ref	0.16 (0.15, 0.17)	Ref	Ref
50 WLM	0.15 (0.14, 0.17)	0.00 (0.00, 0.00)	0.98 (0.96, 0.99)	0.16 (0.15, 0.17)	0.00 (-0.01, 0.00)	0.97 (0.96, 0.99)
20 WLM	0.14 (0.12, 0.15)	-0.02 (-0.03, -0.01)	0.83 (0.78, 0.88)	0.14 (0.13, 0.15)	-0.02 (-0.03, -0.01)	0.83 (0.79, 0.88)
10 WLM	0.12 (0.10, 0.14)	-0.04 (-0.05, -0.02)	0.69 (0.62, 0.77)	0.12 (0.11, 0.14)	-0.04 (-0.05, -0.03)	0.70 (0.63, 0.77)
5 WLM	0.10 (0.08, 0.13)	-0.05 (-0.07, -0.03)	0.59 (0.48, 0.71)	0.10 (0.08, 0.12)	-0.06 (-0.08, -0.04)	0.59 (0.50, 0.71)
2 WLM	0.06 (0.04, 0.08)	-0.10 (-0.12, -0.07)	0.36 (0.25, 0.52)	0.07 (0.04, 0.09)	-0.09 (-0.12, -0.07)	0.37 (0.25, 0.52)
1 WLM	0.07 (0.00, 0.14)	-0.09 (-0.16, -0.01)	0.34 (0.13, 0.93)	0.06 (0.03, 0.08)	-0.11 (-0.14, -0.08)	0.30 (0.18, 0.49)
0.75 WLM	0.04 (-0.01, 0.09)	-0.12 (-0.17, -0.07)	0.20 (0.06, 0.66)	0.05 (0.02, 0.08)	-0.11 (-0.14, -0.08)	0.28 (0.16, 0.48)
0.33 WLM				0.05 (0.02, 0.08)	-0.11 (-0.14, -0.08)	0.26 (0.14, 0.47)
No exposure				0.05 (0.02, 0.08)	-0.12 (-0.15, -0.08)	0.24 (0.12, 0.46)

^a Use of Wald-type confidence intervals resulted in implausible lower confidence intervals at the lowest risk estimate