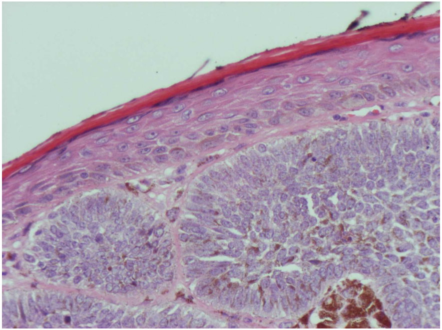
Spatio-Temporal Threshold Models for Relating UV Exposures and Skin Cancer in the Central United States

Laura A. Hatfield and Bradley P. Carlin

Division of Biostatistics
School of Public Health
University of Minnesota

September 14, 2009

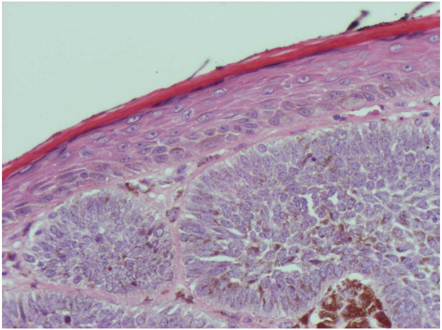
Motivation



Basal Cell Carcinoma, ©DermAtlas

most common form of *non-melanoma skin cancer* (NMSC)

Motivation

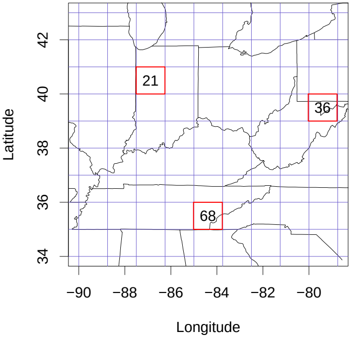


Basal Cell Carcinoma, ©DermAtlas

most common form of *non-melanoma skin cancer* (NMSC)
associated with exposure to ultraviolet (UV) radiation

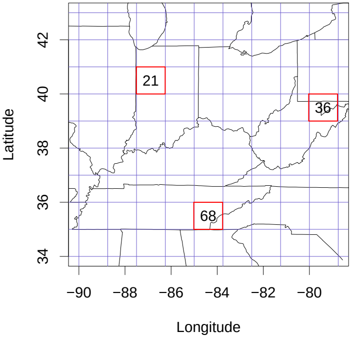
Data Sources

Cancer Outcome Data: from the United States Radiologic Technologists (USRT) cohort



Data Sources

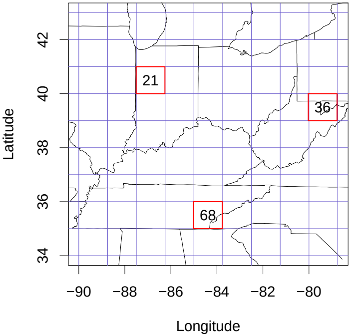
Cancer Outcome Data: from the United States Radiologic Technologists (USRT) cohort



- Cohort: $> 100,000$ US radiation techs

Data Sources

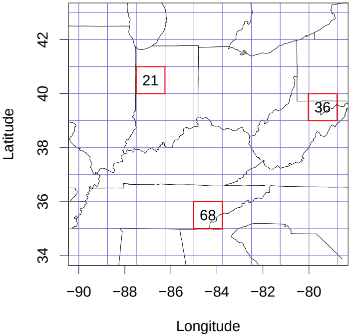
Cancer Outcome Data: from the United States Radiologic Technologists (USRT) cohort



- Cohort: $> 100,000$ US radiation techs
- Our Subsample: $19,081$ living in this square $<$ age 13

Data Sources

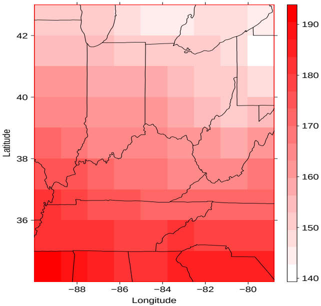
Cancer Outcome Data: from the United States Radiologic Technologists (USRT) cohort



- Cohort: $> 100,000$ US radiation techs
- Our Subsample: $19,081$ living in this square $<$ age 13
- Our Cases: 383 (2.0%) SCC, 1283 (6.7%) BCC, 1495 (7.8%) NMSC (171 both)

Data Sources

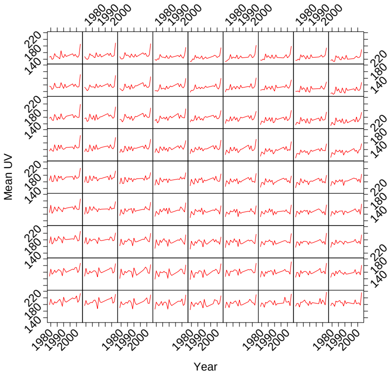
Exposure Data: from the Total Ozone Mapping Spectrometer (TOMS) project

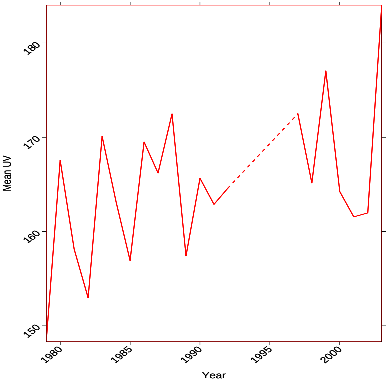


Average Erythemal UV Dose, 1979

UV erythemal dose (milli-watts/m^2)

Data Sources





UV Dose Data by Grid

Enlarged Time Plot of a Single Grid

Note the missing years and the latitude trend; the enlarged grid on the right is the very middle one on the left

Models I: Spatiotemporal UV Dose

We begin by modeling average summer UV dose, Y, in the i^{th} grid and j^{th} year as function of

Models I: Spatiotemporal UV Dose

We begin by modeling average summer UV dose, Y, in the i^{th} grid and j^{th} year as function of

- W_1: Year (1979-2003, centered)
- W_2: Latitude (34.5 to 42.5, centered).

Models I: Spatiotemporal UV Dose

We considered four models:

Models I: Spatiotemporal UV Dose

We considered four models:

- Simple (independent) linear models for year

$$\mu_{ij} = \beta_{0i} + \beta_{1i} W_{1j}$$

Models I: Spatiotemporal UV Dose

We considered four models:

- Simple (independent) linear models for year

$$\mu_{ij} = \beta_{0i} + \beta_{1i} W_{1j}$$

- Random intercepts and time slopes with normal priors

$$\mu_{ij} = \mu_0 + \gamma_{0i} + (\mu_1 + \gamma_{1i}) W_{1j}$$

Models I: Spatiotemporal UV Dose

We considered four models:

- Simple (independent) linear models for year

$$\mu_{ij} = \beta_{0i} + \beta_{1i} W_{1j}$$

- Random intercepts and time slopes with normal priors

$$\mu_{ij} = \mu_0 + \gamma_{0i} + (\mu_1 + \gamma_{1i}) W_{1j}$$

- Same, adjusted for latitude

$$\mu_{ij} = \mu_0 + \gamma_{0i} + (\mu_1 + \gamma_{1i}) W_{1j} + \alpha W_{2i}$$

Models I: Spatiotemporal UV Dose

We considered four models:

- Simple (independent) linear models for year

$$\mu_{ij} = \beta_{0i} + \beta_{1i} W_{1j}$$

- Random intercepts and time slopes with normal priors

$$\mu_{ij} = \mu_0 + \gamma_{0i} + (\mu_1 + \gamma_{1i}) W_{1j}$$

- Same, adjusted for latitude

$$\mu_{ij} = \mu_0 + \gamma_{0i} + (\mu_1 + \gamma_{1i}) W_{1j} + \alpha W_{2i}$$

- Drop random slopes, add spatially-structured prior on random intercepts

$$\mu_{ij} = \mu_0 + \gamma_{0i} + \mu_1 W_{1j} + \alpha W_{2i};\ \gamma_0 \sim CAR(\tau_0)$$

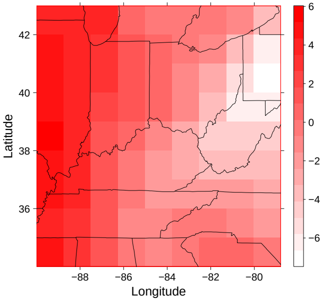
Results I: Spatiotemporal UV Dose

Model	$\overline{D}$	$\hat{D}$	p_D	DIC
Simple linear	715.50	552.50	163.02	878.50
Random effects	650.40	567.30	83.06	733.50
+ Latitude	650.20	582.40	67.81	718.00
+ CAR prior	621.50	579.40	42.04	*663.50*

Model	MSEP (1979)	MSEP (1992)	MSEP (3 grids)
Simple linear	159.51	83.70	NA
Random effects	140.86	80.65	400.00
+ Latitude	145.06	81.27	144.27
+ CAR prior	*140.12*	*79.49*	*121.51*

Model comparison, spatiotemporal UV models.

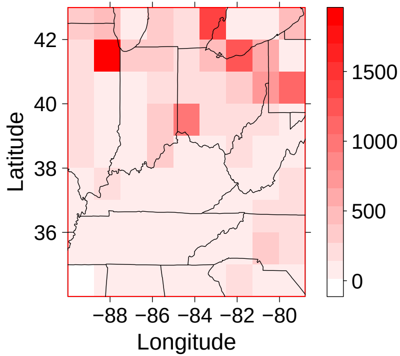
Results I: Spatiotemporal UV Dose



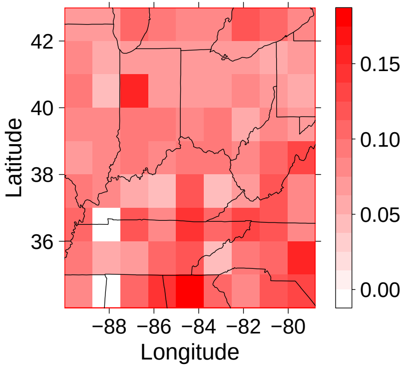
Map of fitted spatially structured random intercepts (γ_{0i})

Parameter	Mean	95% Credible Interval
μ_0 (Intercept)	166.80	[166.40, 167.10]
α (Latitude)	-4.76	[-5.40, -4.19]
μ_1 (Year)	0.59	[0.55, 0.64]
σ_0 (Spatial sd)	3.82	[3.04, 4.75]
σ (Error sd)	7.37	[7.13, 7.62]

Data II: Cancer and Average UV Dose



Cohort Members



NMSC Cumulative Incidence Rate

Models II: Cancer and Average UV Dose

We wish to model the NMSC outcome for the k^{th} person, Z_k, as a function of the following covariates:

- X_1: Age ($\in [43, 94]$, centered)
- X_2: Sex (female=1, male=–1)
- X_3: Number of blistering sunburns ($\in [0, 198]$, centered)
- X_4: Complexion (light=1, other=–1)
- X_5: Average UV dose under age 13
- X_6: Hours spent in summer sun (cumulative under age 13)

Models II: Cancer and Average UV Dose

We wish to model the NMSC outcome for the k^{th} person, Z_k, as a function of the following covariates:

- X_1: Age ($\in [43, 94]$, centered)
- X_2: Sex (female=1, male=–1)
- X_3: Number of blistering sunburns ($\in [0, 198]$, centered)
- X_4: Complexion (light=1, other=–1)
- X_5: Average UV dose under age 13
- X_6: Hours spent in summer sun (cumulative under age 13)

Since the UV exposure trajectories do not cross over time, we may take X_5 as the fitted exposure at any time point for k's grid square

We also considered multiplicative interaction between UV dose and hours in the sun

Models II: Cancer and Average UV Dose

The DIC-best model:

$$
\begin{aligned}
\textit{Stage 1:} \quad & Y_{ij} \sim \textit{Normal}(\mu_{ij}, \sigma^2) \\
& \mu_{ij} = \mu_0 + \gamma_{0i} + \alpha W_{2i} + \mu_1 W_{1j};\ \boldsymbol{\gamma}_0 \overset{iid}{\sim} \textit{CAR}(\tau_0) \\
\textit{Stage 2:} \quad & Z_k \sim \textit{Bernoulli}(p_k) \\
\textit{logit}(p_k) & = \beta_0 + \beta_1 X_{1k} + \beta_2 X_{2k} + \beta_3 X_{3k} + \beta_4 X_{4k} \\
& \quad + \beta_5(\gamma_{0g_k} + \alpha W_{2g_k}) + \beta_6 X_{6k}\ ,
\end{aligned}
$$

where g_k is k's childhood grid square

Results II: Cancer and Average UV Dose

Parameter	Mean	95% Credible Interval
β_0 (Intercept)	-2.473	$[-2.552, -2.396]$
β_1 (Age)	0.054	[0.049, 0.060]
β_2 (Sex)	0.012	$[-0.056, 0.079]$
β_3 (Blistering sunburns)	0.022	[0.017, 0.027]
β_4 (Complexion)	0.328	[0.273, 0.384]
β_5 (Average UV dose)	0.013	[0.008, 0.017]
β_6 (Hours in sun)	0.046	$[-0.010, 0.101]$

Stage 2 posterior parameter estimates

- The UV dose variable has mean 166.3 and sd 15.4, so 1 sd change is .013*15.4 = .20 on the logit scale
- Hours in sun (β_6) is not quite significant, but trend is positive

Models III: Cancer and Peak UV Dose

We now wish to model NMSC outcome (Z) of the k^{th} person as a function of peak UV dose, $X_5(g_k, C)$, i.e., the historical average proportion of summer days $\geq C$ in the k^{th} person's childhood grid square, g_k, where the threshold C is estimated from the data.

We considered the same model variations as above.

Models III: Cancer and Peak UV Dose

We now wish to model NMSC outcome (Z) of the k^{th} person as a function of peak UV dose, $X_5(g_k, C)$, i.e., the historical average proportion of summer days $\geq C$ in the k^{th} person's childhood grid square, g_k, where the threshold C is estimated from the data.

We considered the same model variations as above.

The best model (according to DIC and 10% holdout predictive ability):

$$\begin{aligned} Z_k &\sim Bernoulli(p_k) \\ logit(p_k) &= \beta_0 + \beta_1 X_{1k} + \beta_2 X_{2k} + \beta_3 X_{3k} + \beta_4 X_{4k} \\ &\quad + \beta_5 X_5(g_k, C) + \beta_6 X_{6k} \end{aligned}$$

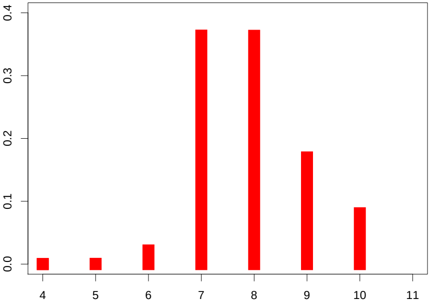
Results III: Cancer and Peak UV Dose

Parameter	Mean	95% Credible Interval
β_0 (Intercept)	-2.475	$[-2.555, -2.397]$
β_1 (Age)	0.054	[0.049, 0.060]
β_2 (Sex)	0.013	$[-0.056, 0.083]$
β_3 (Blistering sunburns)	0.022	[0.017, 0.027]
β_4 (Complexion)	0.329	[0.273, 0.385]
β_5 (Peak UV dose)	1.594	[0.864, 2.847]
β_6 (Hours in sun)	0.046	$[-0.009, 0.101]$
C (Exposure threshold)	6.842	[5, 9]

posterior parameter estimates

- Childhood peak UV dose (β_5) is positively associated with risk
- Hours in sun (β_6) is not significant, but trend is positive

Results III: Cancer and Peak UV Dose



Posterior density of the UV Index threshold (C)

- WHO categories: Low $(0-2)$, Moderate $(3-5)$, High $(6-7)$, Very High $(8-10)$, Extreme $(11+)$
- "Above 7" recommendations: avoid midday sun, seek shade, and use all protective measures (clothing, hat, sunscreen)
- Posterior distribution favors UV Index values in the High to Very High categories for NMSC risk

Data IV: Adding in Ionizing Radiation (IR) Dose

For our USRT cohort, we also have annual occupational ionizing radiation (IR) exposure estimates for each person based on

Data IV: Adding in Ionizing Radiation (IR) Dose

For our USRT cohort, we also have annual occupational ionizing radiation (IR) exposure estimates for each person based on

- Work history: apron usage and procedures performed
- Calendar year: estimates from the literature used
- Measurements: film badges for $\sim$ 30% of person-years ($\sim$ 1970+)

Data IV: Adding in Ionizing Radiation (IR) Dose

For our USRT cohort, we also have annual occupational ionizing radiation (IR) exposure estimates for each person based on

- Work history: apron usage and procedures performed
- Calendar year: estimates from the literature used
- Measurements: film badges for $\sim$ 30% of person-years ($\sim$ 1970+)

Considered 2 versions of IR exposure:

- Lifetime cumulative (until diagnosis or censoring)
- Year-averaged (by number of years worked)

Both strongly related to age ($\rho \approx .5$), via first year worked (for both) and number of years worked (cumulative only).

Note: All of the following results use the entire nationwide USRT cohort of $n = 66,565$ participants

Models IV: UV + IR, Logistic

- Covariates
  - X_1: Age (centered)
  - X_2: Male (1, female$= -1$)
  - X_3: Acute sun sensitivity (1, -1 otherwise)
  - X_4: Light complexion (1, -1 otherwise)
  - X_5: Lifetime blistering burns (centered)
  - X_6: Hours in sun < 13 yrs (standardized)

Models IV: UV + IR, Logistic

- Covariates
  - X_1: Age (centered)
  - X_2: Male (1, female$= -1$)
  - X_3: Acute sun sensitivity (1, -1 otherwise)
  - X_4: Light complexion (1, -1 otherwise)
  - X_5: Lifetime blistering burns (centered)
  - X_6: Hours in sun < 13 yrs (standardized)
- Exposures
  - X_7: UV < 13 yrs (centered)
  - X_8: IR (cumulative or year-averaged, log scale, centered)

Models IV: UV + IR, Logistic

- Covariates
  - X_1: Age (centered)
  - X_2: Male (1, female$= -1$)
  - X_3: Acute sun sensitivity (1, -1 otherwise)
  - X_4: Light complexion (1, -1 otherwise)
  - X_5: Lifetime blistering burns (centered)
  - X_6: Hours in sun < 13 yrs (standardized)
- Exposures
  - X_7: UV < 13 yrs (centered)
  - X_8: IR (cumulative or year-averaged, log scale, centered)
- Model: Logistic again:

$$
\begin{aligned}
Z_k &\sim Bernoulli(p_k) \\
logit(p_k) &= \beta_0 + \beta_1 X_{1k} + \beta_2 X_{2k} + \beta_3 X_{3k} + \beta_4 X_{4k} \\
&\quad + \beta_5 X_{5k} + \beta_6 X_{6k} + \beta_7 X_{7k} + \beta_8 X_{8k}
\end{aligned}
$$

Results IV: UV + IR, Logistic

Age-adjusted models yielded significant effects for UV + IR (using year-averaged IR, but not cumulative IR).

Parameter	Estimate	95% Confidence Interval
β_0 (Intercept)	-3.419	$[-3.59, -3.25]$
β_1 (Age)	0.052	[0.048, 0.055]
β_2 (Male)	0.048	[0.014, 0.082]
β_3 (Acute sun sens.)	0.339	[0.277, 0.400]
β_4 (Complexion)	0.252	[0.221, 0.282]
β_5 (Blistering burns)	0.018	[0.016, 0.021]
β_6 (Hours in sun)	0.041	[0.013, 0.069]
β_7 (UV $<$ age 13)	0.005	[0.004, 0.006]
β_8 (Average IR)	0.050	[0.013, 0.088]

- UV dose variable has sd 29.6, so effect on logit of 1 sd increase is $.00485 * 29.6 = .148$
- IR dose variable has sd 0.835, so effect on logit of 1 sd increase is $0.0504 * .835 = .042$

So UV “has a larger practical effect” on the logit!

Models V: UV + IR, Cox PH

- Two ways to define survival time outcome: time to NMSC or censoring since
  - age 18, or
  - 1996 (approximate date of 2nd survey)

Models V: UV + IR, Cox PH

- Two ways to define survival time outcome: time to NMSC or censoring since
  - age 18, or
  - 1996 (approximate date of 2nd survey)

    Concept: cohort recruited in 1996, followed forward
    (Requires deletion of $n = 2460$ patients diagnosed before 1996)

Models V: UV + IR, Cox PH

- Two ways to define survival time outcome: time to NMSC or censoring since
  - age 18, or
  - 1996 (approximate date of 2nd survey)

    Concept: cohort recruited in 1996, followed forward
    (Requires deletion of $n = 2460$ patients diagnosed before 1996)

    Goals: break connection between age and IR dose, reduce recollection error, and homogenize diagnostic/public health practices

Models V: UV + IR, Cox PH

- Two ways to define survival time outcome: time to NMSC or censoring since
  - age 18, or
  - 1996 (approximate date of 2nd survey)

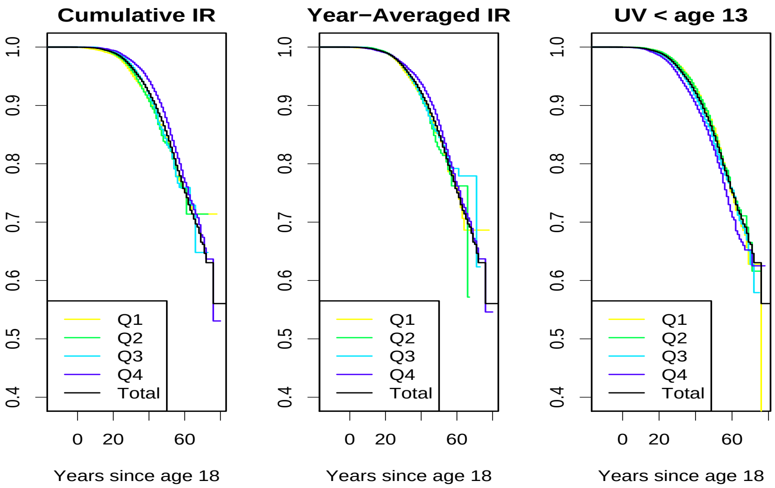
    Concept: cohort recruited in 1996, followed forward
    (Requires deletion of $n = 2460$ patients diagnosed before 1996)

    Goals: break connection between age and IR dose, reduce recollection error, and homogenize diagnostic/public health practices
- Covariates: same as above
- Model: Cox Proportional Hazards (PH)

$$
\begin{aligned}
h(t_k|\mathbf{X}_k) &= h_0(t)\exp(\beta_2 X_{2k} + \beta_3 X_{3k} + \beta_4 X_{4k} \\
&\quad + \beta_5 X_{5k} + \beta_6 X_{6k} + \beta_7 X_{7k} + \beta_8 X_{8k})
\end{aligned}
$$

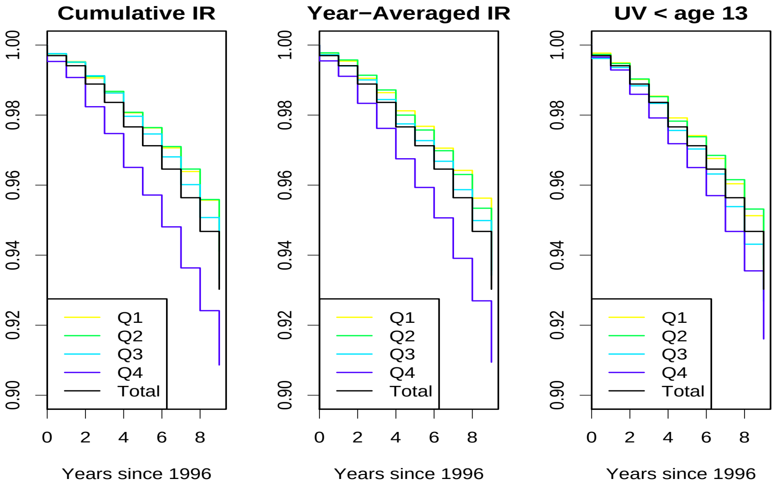
Models V: UV + IR, Cox PH

Age time scale fails to separate the survival patterns by exposure quartile:



Models V: UV + IR, Cox PH

Calendar year time scale succeeds in separating the survival curves by exposure quartile:



Results V: UV + IR, Cox PH

Cohort-type survival framework using calendar year time scale produces sensible (i.e., positive) estimates for IR risk:

Parameter	Hazard Ratio	95% Confidence Interval
β_2 (Male)	1.121	[1.078, 1.166]
β_3 (Acute sun sens.)	1.351	[1.255, 1.454]
β_4 (Light complexion)	1.219	[1.175, 1.265]
β_5 (Blistering burns)	1.013	[1.011, 1.016]
β_6 (Hours in sun < 13)	1.012	[0.978, 1.046]
β_7 (UV $<$ age 13)	1.004	[1.003, 1.005]
β_8 (Average IR)	1.352	[1.294, 1.411]

- UV dose variable has sd 29.6, so Cox HR effect of 1 sd increase is $\exp(0.00375 * 29.6) = 1.12$
- IR dose variable has sd 0.835, so Cox HR effect of 1 sd increase is $\exp(.301 * .835) = 1.29$

So IR "has a larger practical effect" on the HR!

Conclusions so far

Using USRT data from the central US:

- **UV doses increased** over the time period examined, with negligible spatial variation in time trends
- **Spatial variation** in UV dose exists after adjusting for latitude
- **Average childhood UV dose predicts cancer** (NMSC) outcomes *independently* of hours spent in sun
- **Peak childhood UV dose predicts cancer** (NMSC) outcomes best when UV Index threshold is 7 or 8 (i.e., "High to Very High")

Conclusions so far

Using USRT data from the central US:

- **UV doses increased** over the time period examined, with negligible spatial variation in time trends
- **Spatial variation** in UV dose exists after adjusting for latitude
- **Average childhood UV dose predicts cancer** (NMSC) outcomes *independently* of hours spent in sun
- **Peak childhood UV dose predicts cancer** (NMSC) outcomes best when UV Index threshold is 7 or 8 (i.e., "High to Very High")

Using USRT data from the entire US:

- **Childhood UV + IR (avg. lifetime) exposures related to cancer** through logistic models
- **Childhood UV + IR (avg. lifetime) exposures related to cancer** through survival models provided we use a calendar year-based definition for the event time (cohort model)

Future work: Refining UV

- Refinements of the UV dose variable: Letting t indicate *continuous* time within summer, redefine UV exposure as cumulative over ages 3 to 13:

$$UV_k = \frac{1}{150} \sum_{j=1}^{m} \int_0^{150} \mu_{g_k, a_{k+j}}(t) dt \ ,$$

  where g_k indicates subject k's grid of residence at age 10, and a_k indexes the year in which k turned 2 years old

  May combine this additively or multiplicatively with annual (modeled) IR doses, contemporaneously or lagged by L years

- We also have data on the year of diagnosis, facilitating use of possibly more sensitive survival models.

Future work: Handling misalignment

- So far we have ignored spatial misalignment (Zhu et al., *Environmetrics*, 2003) in our data: the UV observations are really the result of a weather-informed model that collects data at a point level but produces grid level exposure estimates, which we assume apply to every resident of the grid
- There is also a form of temporal misalignment: the most relevant (e.g., childhood) UV exposures may vary both spatially and temporally, yet we currently ignore the latter

Future work: Handling misalignment

- So far we have ignored spatial misalignment (Zhu et al., *Environmetrics*, 2003) in our data: the UV observations are really the result of a weather-informed model that collects data at a point level but produces grid level exposure estimates, which we assume apply to every resident of the grid
- There is also a form of temporal misalignment: the most relevant (e.g., childhood) UV exposures may vary both spatially and temporally, yet we currently ignore the latter
- Fuentes et al. (*Biometrics*, 2006) give a generalized Poisson model for areal mortality counts $Y_{jk}(t)$ over strata (e.g., gender, age, etc.) k and counties j by month t
  - Covariates in the log-link include county-level confounders (e.g., ozone) $\mathbf{U}_j(t)$, strata-level confounders (e.g., gender, age, etc.) $\mathbf{X}_k(t)$, and I different PM exposure species $\mathbf{V}_j(t)$.
  - The spatial misalignment is resolved by integrating an underlying "total" PM surface $Z(s,t)$ over the counties C_j, where this is in turn modeled linearly as a function of weather variables $W(s,t)$.

Fuentes et al. (2006) model

For month t, county j, and stratum k,

$$
\begin{aligned}
Y_{jk}(t) &\sim GPoi(\alpha, \mu_{jk}(t)) \\
log(\mu_{jk}(t)) &= \mathbf{U}_j^T(t)\gamma_1 + \mathbf{X}_k^T(t)\gamma_2 + \mathbf{V}_j^T(t)\beta, \\
\text{where} \quad & \mathbf{V}_j^T(t) = (V_{1j}(t), \ldots, V_{Ij}(t)) \\
V_{ij}(t) &= \xi_i(t) \int_{C_j} Z(s,t)ds + \phi_{ij}(t) \\
\phi_j(t)|\phi_{-j}(t) &\sim N\left(\sum_{j' \neq j} B_{jj'}(t)\phi_{j'}(t), \frac{1}{m_j}\Sigma_\phi(t)\right), \\
\text{where} \quad & \phi_j(t)^T = (\phi_{1j}(t), \ldots, \phi_{Ij}(t)) \\
Z(s,t) &= \mathbf{W}^T(s,t)\boldsymbol{\eta} + \epsilon(s,t)
\end{aligned}
$$

Fuentes et al. (2006) model

For month t, county j, and stratum k,

$$
\begin{aligned}
Y_{jk}(t) &\sim GPoi(\alpha, \mu_{jk}(t)) \\
log(\mu_{jk}(t)) &= \mathbf{U}_j^T(t)\gamma_1 + \mathbf{X}_k^T(t)\gamma_2 + \mathbf{V}_j^T(t)\boldsymbol{\beta}, \\
\text{where} \quad & \mathbf{V}_j^T(t) = (V_{1j}(t), \ldots, V_{Ij}(t)) \\
V_{ij}(t) &= \xi_i(t) \int_{C_j} Z(s,t)ds + \phi_{ij}(t) \\
\phi_j(t)|\phi_{-j}(t) &\sim N\left(\sum_{j' \neq j} B_{jj'}(t)\phi_{j'}(t), \frac{1}{m_j}\Sigma_\phi(t)\right), \\
\text{where} \quad & \phi_j(t)^T = (\phi_{1j}(t), \ldots, \phi_{Ij}(t)) \\
Z(s,t) &= \mathbf{W}^T(s,t)\boldsymbol{\eta} + \epsilon(s,t)
\end{aligned}
$$

- Here, $\xi_i(t)$ represents the proportion of total exposure at time t explained by PM species i
- a multivariate CAR (MCAR) model for the $\phi_{ij}(t)$ allows covariance structure $\Sigma_\phi(t)$ that evolves over time

Future Work: Gaussian Process model

Returning to our setting and again thinking of time within summer t as continuous, modify our UV model for grid i to

$$Y_i(t) = \mu_i(t) + \epsilon_i(t) \ ,$$

where the $\epsilon_i(t)$ are from a white noise process, and

$$\mu_i(t) = \mu_0 + \gamma_{0i} + \alpha W_{2i} + Z_i(t) \ .$$

Future Work: Gaussian Process model

Returning to our setting and again thinking of time within summer t as continuous, modify our UV model for grid i to

$$Y_i(t) = \mu_i(t) + \epsilon_i(t) \ ,$$

where the $\epsilon_i(t)$ are from a white noise process, and

$$\mu_i(t) = \mu_0 + \gamma_{0i} + \alpha W_{2i} + Z_i(t) \ .$$

Writing

$$\mathbf{Z} = (\mathbf{Z}'(t_1), \ldots, \mathbf{Z}'(t_T))',$$

where

$$\mathbf{Z}(t) = (Z_1(t), \ldots, Z_{81}(t))'$$

for any t, assume the $Z_i(t)$ follow a spatial Gaussian process (GP) model...

Future Work: Gaussian Process model

Specifically, model $\mathbf{Z} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ where

$$\boldsymbol{\Sigma} = R(\phi) \otimes \tau^2 (D - \lambda W)^{-1} ,$$

the Kronecker product of

- $T \times T$ temporal correlation matrix $R(\phi)$ (say, (t_1, t_2) element $\exp(-\phi|t_1 - t_2|)$), and
- $M \times M$ spatial correlation matrix (say, of the usual CAR type)
  - Adjacency matrix W, with $w_{ii'} = I(i' \in \partial_i)$, where ∂_i is the neighbor set of i
  - D is diagonal with $D_{ii} = w_{i+}$
  - Smoothing parameter $\lambda \in (0, 1)$

Future Work: Gaussian Process model

Specifically, model $\mathbf{Z} \sim N(\mathbf{0}, \mathbf{\Sigma})$ where

$$\mathbf{\Sigma} = R(\phi) \otimes \tau^2 (D - \lambda W)^{-1} \ ,$$

the Kronecker product of

- $T \times T$ temporal correlation matrix $R(\phi)$ (say, (t_1, t_2) element $\exp(-\phi|t_1 - t_2|)$), and
- $M \times M$ spatial correlation matrix (say, of the usual CAR type)
  - Adjacency matrix W, with $w_{ii'} = I(i' \in \partial_i)$, where ∂_i is the neighbor set of i
  - D is diagonal with $D_{ii} = w_{i+}$
  - Smoothing parameter $\lambda \in (0, 1)$
- We are interested in the temporal gradient, $Z_i'(t)$:
  - We seek the t where this becomes significantly > 0
  - May vary both temporally (function of t) and spatially (for each i)

Final thoughts on the spatial GP model

- Note that while the derivative process is never observed, inference may still proceed from the predictive distribution, $P(Z_i'(t_0)|\mathbf{Z})$
- Possible to compute because:
  - covariance functions yield mean square differentiable paths, and
  - sampling from the posterior predictive easy via composition (Gaussian kernel with the Gibbs posterior draws)

Final thoughts on the spatial GP model

- Note that while the derivative process is never observed, inference may still proceed from the predictive distribution, $P(Z_i'(t_0)|\mathbf{Z})$
- Possible to compute because:
  - covariance functions yield mean square differentiable paths, and
  - sampling from the posterior predictive easy via composition (Gaussian kernel with the Gibbs posterior draws)
- Potentially important in our IR/UV/NMSC setting since if UV exposure is increasing over time, identical self-reports of exposure from widely different years may correspond to quite different cancer risks!
- Ideally, all of this structure would be added to our logistic model for the NMSC responses, now indexed by both individual k and exposure epoch t.

THE END (for now)