The Nevada Test Site Fallout Data

Raymond J. Carroll
Department of Statistics
Institute of Applied Mathematics and Computational Science
Texas A&M University

Joint work with Annamaria Guolo, Owen Hoffman and Yehua Li

September 11, 2011

Owen Hoffman

Yehua Li

Annamaria Guolo

Personal History

- I was born in **Yokohama Japan**
- Since I am not native born, I cannot be **President of the United States**
- We lived in Virginia, Nebraska, Germany and **Texas**
- I graduated from the University of Texas at Austin
- My Ph.D. is from Purdue University

Personal History

- From 1974-1987, I was a professor at the **University of North Carolina**
- Since 1987, I have been a professor of Statistics, Nutrition and Toxicology at **Texas A&M University**
- I spend much time at **National Cancer Institute**, in the same division as Andre Bouville and Mark Little
- I love living in Texas, and am very proud to be a **Texan**

West Texas ☺

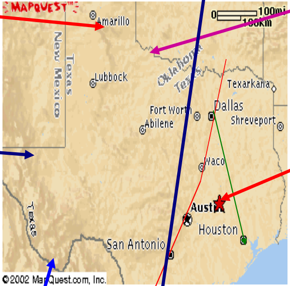
East Texas ☹

Palo Duro Canyon, the Grand Canyon of Texas

Wichita Falls, Wichita Falls, that's my hometown

Guadalupe Mountains National Park

College Station, home of Texas A&M University



I-45

Big Bend National Park

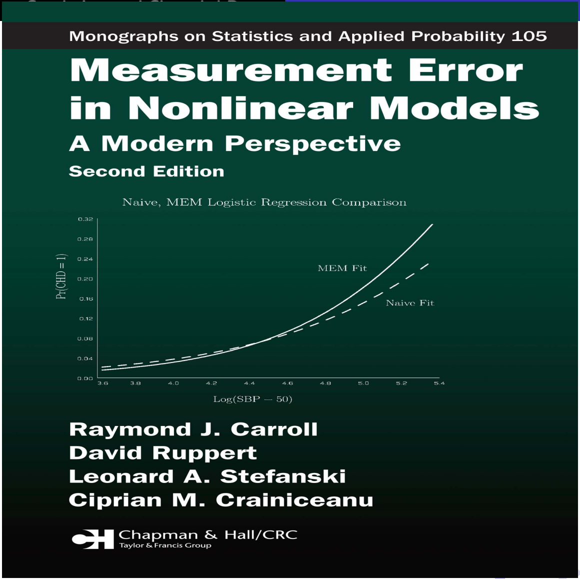
I-35

Palo Duro Canyon, Amarillo Texas

Palo Duro Canyon of the Red River

Research

- My primary interest is in **Measurement Error Models**
- Major applications are in **Nutritional Epidemiology**, because usual diet ascertainment is very difficult.
- A second application is **Radiation Epidemiology**, because of the uncertainty in dose ascertainment.


Monographs on Statistics and Applied Probability 105
Measurement Error
in Nonlinear Models
A Modern Perspective
Second Edition
Naive, MEM Logistic Regression Comparison
Pr(CHD = 1)
0.32
0.28
0.24
0.20
0.16
0.12
0.08
0.04
0.00
MEM Fit
Naive Fit
3.6
3.8
4.0
4.2
4.4
4.6
4.8
5.0
5.2
5.4
Log(SBP − 50)
Raymond J. Carroll
David Ruppert
Leonard A. Stefanski
Ciprian M. Crainiceanu
Chapman & Hall/CRC
Taylor & Francis Group

Administration

- I direct the **Institute for Applied Mathematics and Computational Science**
- The members are from Statistics, Applied Mathematics, Computer Science, Nuclear Engineering, and many other fields
- It is financially supported by KAUST, the King Abdullah University of Science and Technology
- KAUST is the first western-style university in Saudi Arabia

KAUST Map

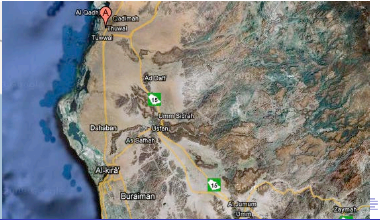
TEXAS A&M UNIVERSITY
INSTITUTE FOR APPLIED MATHEMATICS AND COMPUTATIONAL SCIENCE

KAUST

Campus

Approximately 20 square miles

Located in Thuwal (50 miles north of Jeddah)



Acknowledgments

- I gratefully acknowledge research support from the **National Cancer Institute** (R37-CA057030)
- My research is also supported by **KAUST** (Award Number KUS-CI-016-04)

Outline

- Introduction to Measurement Error Models
- Nevada Test Site
- Model and Methods
- Data Analysis
- Conclusions and Chernobyl Data

Measurement Error Models

- Measurement error problems refer to the regression problems in which one or more covariate is measured with uncertainty.
- This can be due to systematic bias and random error.
- The classic case is of course radiation dose absorbed by the thyroid

Measurement Error Models

- I will always let Y denote the response
- The true dose will be X, but it is not observable
- The calculated dose is W.
- A **crucial assumption** if that W is a **surrogate**, i.e., it carries no extra information about Y if X were observable.

Effects of Measurement Errors

- Effects of measurement errors:
  - **Bias** in parameter estimation.
  - **Loss of statistical power** to detect health risks
  - **Loss of Subtle Features**, e.g., threshold effects cannot be seen in the observed data

Classical Error

- **Classical Measurement Error** in radiation dosimetry occurs when $\log(W) = \log(X) + U$, with U having mean zero. Thus, **unbiased in the log scale**
- This assumes that calculated dose is dependent only on direct information obtained from an individual.
- It has the property that $\text{var}\{\log(W)\} > \text{var}\{\log(X)\}$.
- The effect of classical error is well-known: usually the **relative risks are attenuated**, i.e., are closer to the null value.

Classical Error

- $\log(W) = \log(X) + U$, where U is a zero-mean random error independent with X.
- The classical model is plausible if W is a direct measurement of X. For example, in the **Nevada Test Site** study, mothers were asked to tell about milk intake and milk source for their children

Berkson Error

- In the **Berkson Error Model**, $\log(X) = \log(W) + U$, where U is a zero-mean random error independent of W.
- In the Berkson model, $\text{var}\{\log(X)\} > \text{var}\{\log(W)\}$.
- In radiation dosimetry, the Berkson model occurs regularly
- At some point, all people sharing certain characteristics (gender, age, location) are assigned the same dose.

Berkson Error

- In the **Berkson Error Model**, $\log(X) = \log(W) + U$
- Thus, true log-dose varies around the calculated log-dose
- In radiation dosimetry, **there are both Berkson and Classical Errors**!

Chernobyl

- In Kukush, et al. (2011), the classical errors arise in the measurement of ^{131}I
- The Berkson errors arise in the measurement of thyroid mass.

Alexander Kukush

Nevada Test Site

- The United States did above-ground nuclear testing in the 1950's, exploding the devices in Area 51 in Nevada
- At least twice, they did so when the winds were strong and blowing east towards Utah and eastern Nevada.
- Outside of Salt Lake city, the affected area was not heavily populated, and was quite rural.
- Many of the children got their milk from their own cows

Original Study

- The original study was led by **Lynn Lyons**, with original dosimetry done by **Andre Bouville**
- After the original paper was published, **Owen Hoffman** joined Andre and performed a massive exercise to redo the dosimetry
- It is this data that I will discuss
- It was planned that there would be much greater followup of the patients for evidence of thyroid disease, but the CDC removed funding for the project
- Some think this was partly due to fears of litigation

Details

- The study had 2,491 individuals exposed to radiation as children
- In the original followup data, there were 20 cases of thyroid neoplasm
- There were over 123 cases of thyroiditis
- Simple analysis based on calculated doses showed a statistically significant effect for thyroiditis
- The effect was suggestive but not significant for neoplasm

Goal

- We set out to try to understand the effects of measurement error of doses on relative risks and statistical power.
- The difficulty is to estimate the radiation dose that they were exposed to 50 years ago.

How were people exposed to radiation?

- Simplistically, radioactive iodine generated from the nuclear test fell on the ground, transmitted into vegetables and cows who ate the grass on the land.
- People were exposed to radiation by consuming vegetables and milk from these cows.

Reconstruction of the dosimetry

The radiation dose for an individual was calculated with a mathematical model taking into account of the following factors:

- Age at exposure
- Gender
- Residence history
- x-ray history
- Whether the individual was breast-fed during childhood
- A diet questionnaire filled out by the parents for milk and vegetables

Berkson and Classical Error

- It was thought that both Berkson and Classical errors exist in the dosimetry
- The **classical errors** occur from the questionnaires filled out by the parents
- The **Berkson errors** occur because of a radiation transport model, so that everyone who had certain measured characteristics was assigned the same calculated dose.

Uncertainties

- Simulations and other calculations were done to estimate the total uncertainty in the calculated doses.
- The data file contained the geometric standard deviation of the total uncertainty, for each individual.
- In the log-dose scale, we have the standard deviation of total uncertainty for each individual.
- The grant was going to redo this aspect. However, in our case, **we do not know what percentage of the uncertainty is Berkson, and what percentage is classical**

Error Structure

- The measurement error is a mixture of Berkson and classical error.
- W is the calculated log-dose, X is the true log-dose, then one way to have both Berkson and classical errors is the model

$$W = L + U_c, \quad X = L + U_b,$$

where L is a latent variable, U_c and U_b are the classical error and the Berkson error respectively, and they are independent of L.

Error Structure

- Model

$$\log(W) = L + U_c, \quad \log(X) = L + U_b$$

- Here L is a latent variable, U_c and U_b are the classical error and the Berkson error respectively, and they are independent of L.
- The total uncertainty $\text{var}(U_c) + \text{var}(U_b)$ is known, and varies from individual to individual
- However, neither variance is known.

Strata Structure

- To conduct a better analysis, we also classify all individuals into different strata according to
  - Type of exposure, i.e. the nuclear tests called Shot Harry, Shot Smoky and all the others.
  - Location of the exposure, e.g. Washington County of Utah, Lincoln County of Nevada.
  - Source of milk the individual drank, e.g. from their own backyard or from a local commercial supply.
- We have a total of 36 strata. The size of a strata varies from 3 to 274.

Shared Berkson Error

- In the NTS study, part of the Berkson error is shared within each stratum, and can be viewed as a stratum-level random effect.
- The Berkson errors are correlated within each stratum, with a compound symmetry structure.
- The amount of correlation varies between different strata.

Expert Advice

- We do not know what percentage of the total uncertainty is Berkson
- We do not know the correlation of the Berkson uncertainty for individuals in any given strata.
- In both cases, Owen Hoffman gave us ranges for the possible values.

The Data

- We have the following information
  - Y_{is} is the binary disease status for the i^{th} subject in the s^{th} stratum. We consider two diseases, thyroiditis and neoplasm, and analyze them separately.
  - W_{is} is the calculated dose.
  - X_{is} is the true dose.
  - Z_{is} is a vector including 1 and gender.
  - $\sigma^2_{is,total}$ estimated total uncertainty in log-dose.

Model

- Excess Relative Risk Model

$$\mathrm{pr}(Y_{is} = 1 | Z_{is}, X_{is}) = H\left\{Z_{is}^{\mathrm{T}}\beta + \log\left(1 + \theta X_{is}\right)\right\}. \qquad (1)$$

- $H(x) = 1/\{1 + \exp(-x)\}$ is the logistic function
- θ is the excess relative risk per gray.

Model

- Mixture of Classical and Berkson error,

$$\log(X)_{is} = L_{is} + U_{is,\mathrm{B}}, \quad \log(W)_{is} = L_{is} + U_{is,\mathrm{C}}$$

- All random variables are normally distributed
- The L_{is} have mean and variance depending on strata and gender
- The Berkson errors $U_{is,\mathrm{B}}$ are correlated within stratum
- We know $\mathrm{var}(U_{is,\mathrm{B}}) + \mathrm{var}(U_{is,\mathrm{C}})$.

Model

- $\sigma^2_{is,B} + \sigma^2_{is,C} = \sigma^2_{is,total}$.
- $\sigma^2_{is,C} = \gamma_C \sigma^2_{is,total}$, where γ_C is the proportion of classical error which is the same for all individuals.
- Shared Berkson Error: $\text{corr}(U_{i_1 s,B}, U_{i_2 s,B}) = \rho_s$, which is different for different strata.
- Ranges for γ_C and the ρ_s were given to us by Dr. Hoffman.

Bayes Prior Distributions

- We used Bayesian and Monte-Carlo EM calculations to give 2 looks to the data.
- The terms γ_C and the ρ_s were uniformly distributed on their ranges.

MCMC

- In the full model we considered, we need to update β, θ, γ_C μ_s, σ_s and ρ_s, which are over 100 parameters.
- We also need to update 2 latent variables X_{is} and L_{is} for every individual. That is nearly 5,000 latent variables.
- We got 100,000 posterior samples, and threw out the first 10,000 samples as burn-in.
- Bayesian inference is automatic: we use the posterior mean as point estimator for the parameters, and the posterior percentiles to construct credible intervals.

Why Bayes?

- The Bayesian computation is computationally demanding
- We needed it however to allow for the uncertainty about the % of uncertainty that is Bayesian, and the correlation of Berkson errors within each stratum
- If we knew these quantities, then much simpler methods are available, such as **regression calibration**, which replaces true dose by its best approximation.

The Frequentist Method – MCEM

- For those who do not like Bayesian calculations, we also implemented a Monte-Carlo EM algorithm
- Very technical, very slow computationally, and not as flexible.
- Also, harder to do inference, such as confidence intervals for excess relative risk per grey.

Back to the NTS data

- The subjects are the 2491 people who were exposed to nuclear radiation during childhood.
- 123 cases of thyroiditis.
- 20 cases of thyroid neoplasm.

Thyroiditis results

Model	Estimate or Posterior Mean	Posterior Median	95% C.I.	
BC Mixture shared	9.35	8.60	3.11	19.39
BC Mixture unshared	9.10	8.47	3.37	18.22
Berkson shared	6.42	5.91	2.15	13.28
Berkson unshared	6.49	6.05	2.55	12.61

Table: Thyroiditis data: inference for the excess relative risk per gray (θ).

Posterior distribution for Thyroiditis excess relative risk

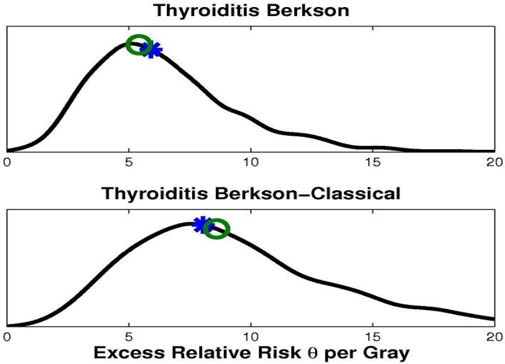


Figure: The circle indicates the posterior median.

Neoplasm results

Model	Estimate or Posterior Mean	Posterior Median	95% C.I.	
BC Mixture shared	60.85	37.93	5.39	240.32
BC Mixture unshared	60.94	38.83	6.04	243.26
Berkson shared	40.73	19.77	3.21	224.66
Berkson unshared	35.17	19.51	3.41	187.81

Table: Neoplasm data: inference for the excess relative risk per gray (θ).

Posterior distribution for Neoplasm excess relative risk

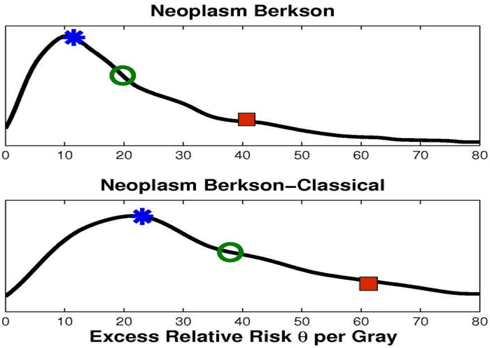


Figure: The circle indicates the posterior median and the square indicates the posterior mean.

Kukush, et al. (2011)

- From Kukush, et al. (2011, *International Journal of Biostatistics*)
- Calculated dose

$$D_i^{\text{calc}} = f_i C_i^{\text{meas}} / M_i^{\text{meas}}$$

- Here C_i^{meas} = the measured content of ^{131}I
- M_i^{meas} is the estimated mass of the thyroid
- f_i = takes into account the parameter values of an ecology-metabolism model

The Classical Part

- The Classical error arises in the measurement of ^{131}I, via the model

$$C_i^{\mathrm{meas}} = C_i^{\mathrm{true}} V_i^C .$$

The Berkson Part

- The Berkson error arises in the measurement of thyroid mass, via the model

$$M_i^{\mathrm{true}} = M_i^{\mathrm{meas}} V_i^M$$

The NTS and Chernobyl Models

- The model in Kukush, et al. is mathematically the same as that of the Nevada Test Site

$$
\begin{aligned}
L_i &= \log(C_i^{\mathrm{true}}) - \log(M_i^{\mathrm{meas}}) + \log(f_i) \\
X_i &= L_i + \log(V_i^M); \\
W_i &= L_i + \log(V_i^C)
\end{aligned}
$$

The NTS and Chernobyl Models

- The huge difference in the analysis is that in Kukush, et al., the Berkson error variance and the classical error variance "have been reliably estimated (see Likhtarev et al., 1993)".
- A second difference is that there appears to be no need for shared Berkson correlation
- These are massive, important differences that allow for much more straightforward analysis

Conclusions

- The NTS study and the Chernoby study are examples where the measurement error structure consists of Berkson and Classical components
- Both studies can be modeled in the same latent variable framework

- The NTS study has much less information available, especially about the relative sizes of the Berkson and Classical error variances.

- Consequently, to account for the additional uncertainties, appropriate modeling is much more complex.