THE RELATIONSHIP BETWEEN CARCINOGEN-INDUCED DNA ADDUCT LEVELS IN DISTAL AND PROXIMAL REGIONS OF THE COLON

Jeffrey S. Morris

M.D. Anderson

Naisyin Wang & **Raymond J. Carroll**

Texas A&M University

Acknowledgements to Joanne R. Lupton, Robert S. Chapkin, Nancy D. Turner and Meeyoung Hong, **Human Nutrition**, TAMU

OUTLINE

- Scientific Background
- Statement of Problem
- General model
- Parametric approach/results
  - Surprising results
  - Robustness checks needed
- Nonparametric approach/result
- Bayesian approach – based on true cell positions
- Conclusions

SCIENTIFIC BACKGROUND: DNA ADDUCT LEVEL

- General Goal: To link dietary variables with factors present during early colon carcinogenesis.
- *DNA adduct level* : amount of certain type of DNA cell damage.
  - Occurs after exposure to particular carcinogen (AOM).
- If damage not repaired or removed, tumor may result.
- We have of course worked on apoptosis and repair.

SCIENTIFIC BACKGROUND: STRUCTURE OF COLON

- **Proximal** region front of colon; **Distal** region back.
- Colon cells replicate and grow within discrete units called ”crypts”.
- We are interested in behavior of cells accounting for their depth in the crypt
  - Relative cell position used, called X (relative depth within crypt).
  - Measures what fraction of the crypt from the bottom a cell is.
  - $X = 0$: bottom of crypt, $X = 1$: top of crypt.
  - $X = 1/2$: middle of crypt

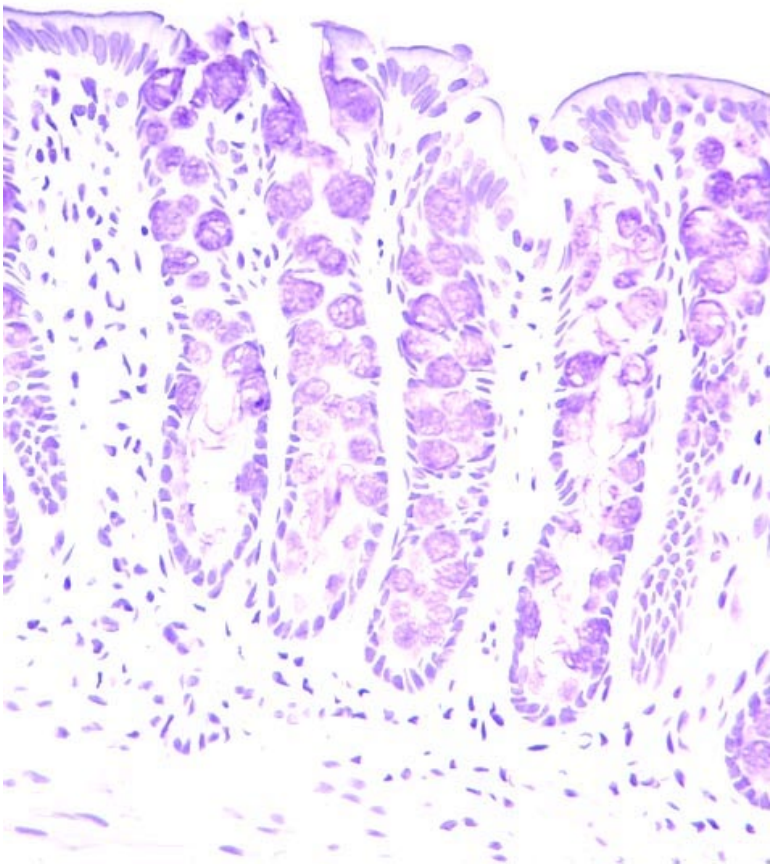

Figure 1: **Cross-sectional view of some colonic crypts**

SCIENTIFIC BACKGROUND: STRUCTURE OF COLON

SCIENTIFIC BACKGROUND: DESCRIPTION OF AOM DATA

- 30 rats randomized to **fish oil** or **corn oil** diets.
- 3 rats/diet controls, other rats exposed to AOM.
  - Euthanized after 3,6,9, or 12 hours.
- In both distal and proximal regions, **$\sim$20 crypts/rat** selected
- Measurement taken for cells along left wall of each crypt
  - Roughly **40 cells/crypt**

SCIENTIFIC BACKGROUND: DIET LINKS WITH COLON CANCER

- Studies indicate diet/colon cancer link.
  - **Fish-eating societies**: lower colon cancer frequencies
  - High corn oil less protective than high fish oil diet.
- **AOM study**: DNA Adduct levels ↑ for corn oil rats?
  - Yes, in distal region.
- **Question**: What is the relationship of adduct levels in distal and proximal regions **within the same rat**?

STATEMENT OF PROBLEM

- *Question*: Is there a relationship between distal and proximal DNA adduct levels within same rat?
  - Specifically, for rats with ↑ distal, also ↑ **proximal**?
  - Relationship depend on **diet** or **relative cell position** (X)?
- Statistically, estimate **correlation function** of two responses as **function of longitudinal covariate**.
  - Two responses: DNA adduct levels in distal and proximal regions (Y_d and Y_p)
  - Covariate: Relative cell position X.
- Note: Structure of data makes this problem nontrivial.

General Model

- Model for distal and proximal measurements as function of X:

$$Y_{jrc_d}(X) = m_j(X,d) + m_{jr}(X,d) + m_{jrc_d}(X,d) + \epsilon_{jrc_d}(X,d$$
$$Y_{jrc_p}(X) = m_j(X,p) + m_{jr}(X,p) + m_{jrc_p}(X,p) + \epsilon_{jrc_p}(X,p$$

  - m_j: fixed mean functions, time j;
  - m_{jr}: rat-level random effect functions, rat r;
  - m_{jrc}: crypt-level random effect functions, crypt c;

- Fact: Y_d and Y_p never observed on same crypts or cells.

- **Correlation of Interest**: $\rho(X)$
  - Between $m_{jr}(X,d)$ and $m_{jr}(X,p)$.

PARAMETRIC MODELING APPROACH

- Assume **parametric structure** for fixed and random functions
- Lots of data analysis and model fitting went into choice of the final model.
  - Has quadratic cell position effects
- Standard Gaussian linear mixed model.
- $\rho(X)$ is a function of rat-level covariance matrices.
- **PROC MIXED in SAS** used for calculations.
- **Parametric bootstrap** used for inference
  - MIXED standard errors suspicious for correlations

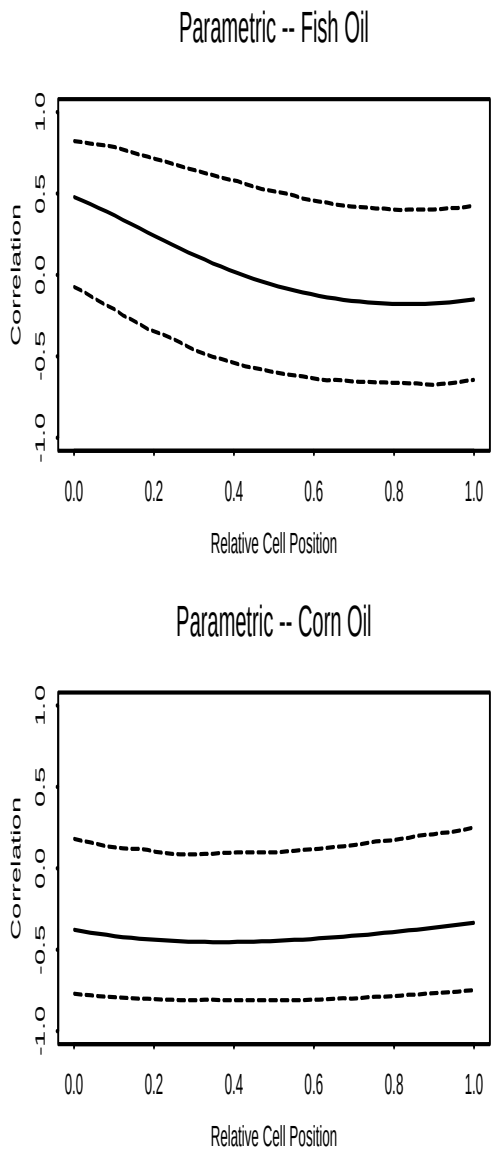

Figure 2: **Parametric estimates for with 90% confidence bounds**

NONPARAMETRIC MODELING: SUMMARY

- The results are **surprising**
- We set out to verify the results by alternative fitting methods
- First method uses **Functional Data Analysis** approach

- No parametric structure assumed on functions.
- Summary of method:
  - **Estimate crypt-level functions** on fixed grid.
  - **Fit mixed model** at each grid point.
  - **Estimated correlation** computed from covariances estimates.
  - Completely nonparametric

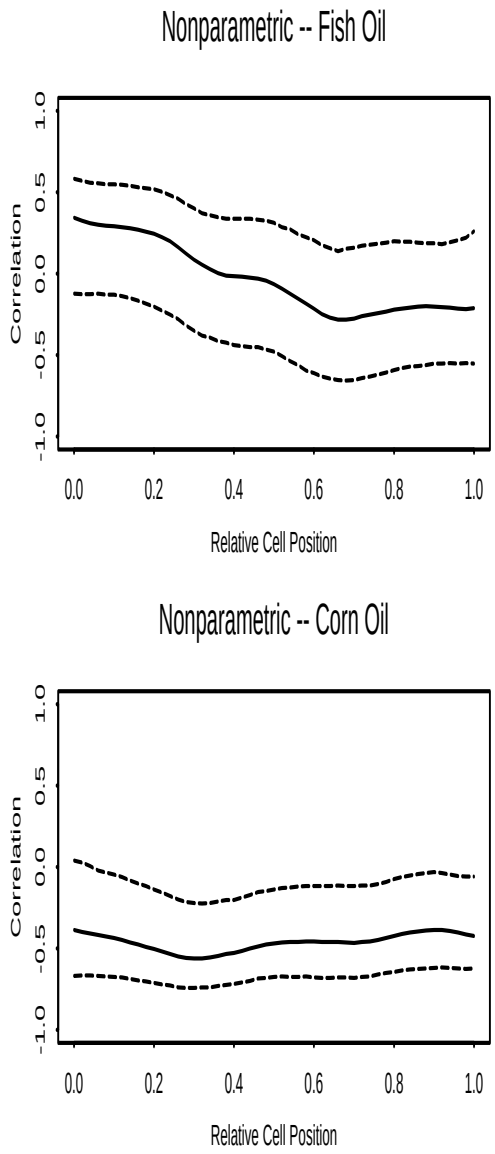

Figure 3: **Nonparametric estimates for with 90% confidence bounds**

CELL POSITION PROBLEM: INTRODUCTION

- **Relative cell position** measure computed from nominal, not true cell positions.
  - **True cell position** tedious to measure.
  - Identical if cells uniform size, shape, and distribution.
- **True cell position** may be more scientifically relevant.
  - Nominal and true scales may yield **different correlations**.
  - Negative corn correlation may be due to nominal scale.
- Internal validation data set
  - True cell position measured for 3 crypts/region/rat.

CELL POSITION PROBLEM: UNIFORMITY

- Cell positions within given crypt modeled as order statistics of Beta(a, b)
- Question: **Is distribution uniform** (i.e. $a = b = 1$)?
- **Bayesian**: Metropolis-Hastings algorithm used to sample posterior
- Results:
  - Proximal region – uniform model OK, **relative cell positions fine to use**.
  - Distal – Not uniform – **more cells at bottom of crypts**
- Correlation based on true cell positions could be different, **in principle**.

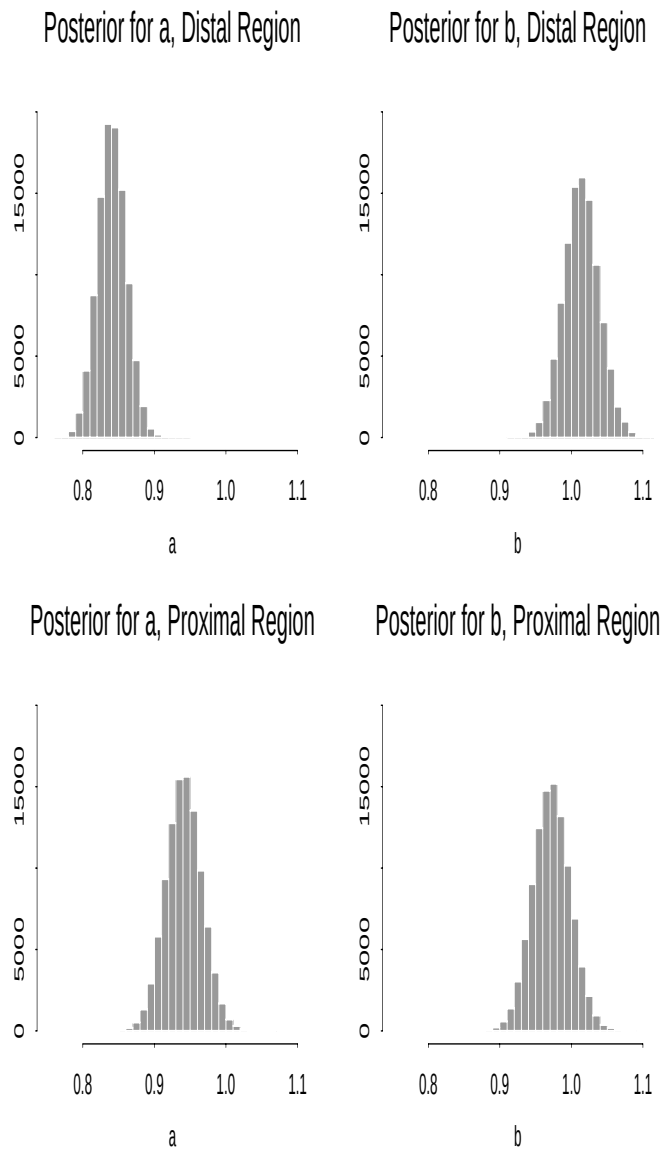

Figure 4: **Posterior distribution of Beta parameters, distal and proximal regions**

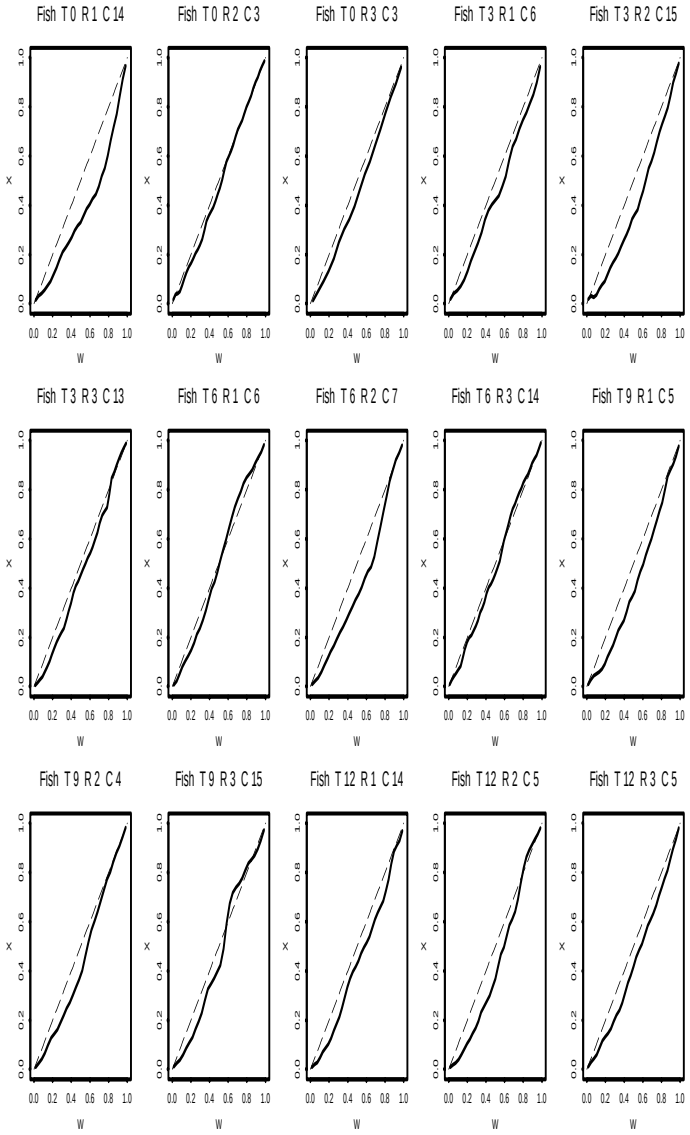

Figure 5: **Portion of distal validation data. Note the consistent pattern.**

CELL POSITION PROBLEM: CORRELATION FUNCTION

- Goal: Estimate **corn correlation function** based on **true measure**.
- Hierarchical linear mixed model with covariates missing for many crypts (about 70%)
  - **Bayesian approach** – Gibbs sampler with two M-H steps.
  - One M-H step updates missing cell positions
  - other updates cell position Beta parameters (a, b).
- Results same based on either true or nominal measure.
  - **Final answer: Negative correlation** not result of using nominal scale.

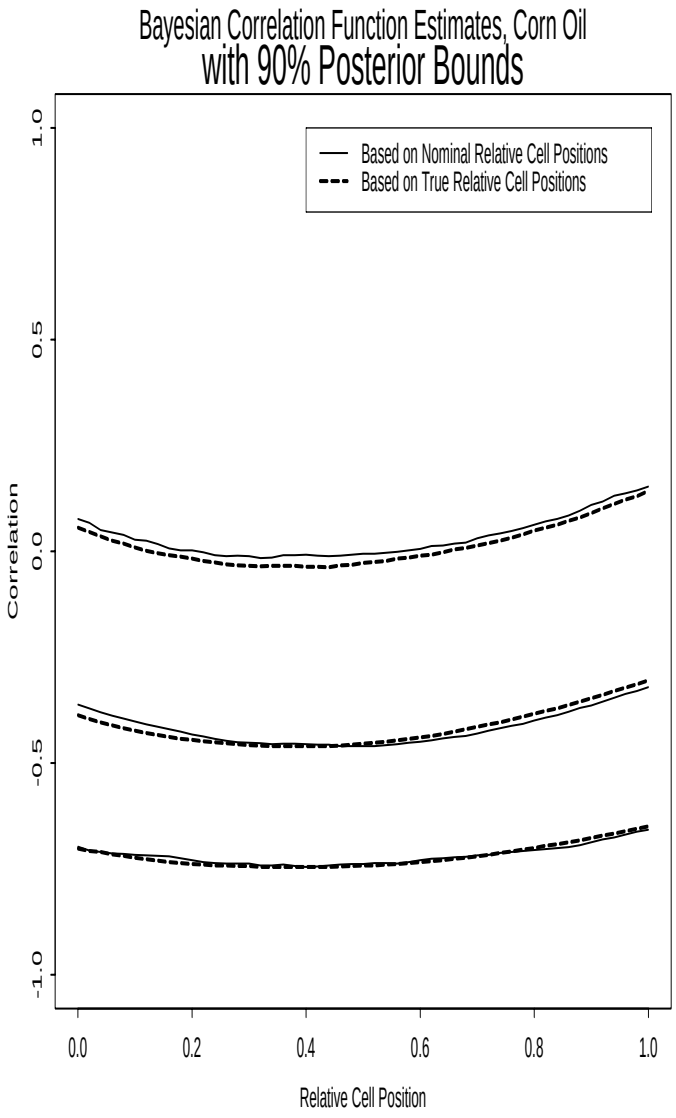

Figure 6: **Corn oil correlation function based on nominal and true scales.**

RESULTS/SCIENTIFIC IMPLICATIONS

- We have fit various models, all results similar
- Summary of Results:
  - Fish oil rats: $\rho > 0$ at bottom of crypts.
  - Corn oil rats: $\rho < 0$ throughout all depths of crypts.
  - Diet difference **statistically significant** near bottom of crypts.
- Implications:
  - Localization of cell damage for corn oil rats.
  - More designed studies needed.