Statistical Methods for Behavioral Economics Data in Tobacco Research

Xianghua Luo

Division of Biostatistics, School of Public Health
University of Minnesota

March 15, 2022

Outline

1. Background
2. Survival Analysis Methods Research
   - Recurrent event gap time methods
   - Intermittently observed time-dependent covariates
   - Composite endpoints with component-wise censoring
   - Graphical methods for survival endpoints in clinical trials
   - Collaborators
3. Statistical Methods for Behavioral Economics Data in Tobacco Research
   - Cigarette Purchase Task (CPT) Data
   - Left-censored mixed effects model
   - Two-part mixed effects model
   - Bayesian hierarchical model for meta-analysis
   - Collaborators and reference

Background

Education:

1997	**B.S.**, Geography, Peking University, China
	Minor, Computer Science, Peking University, China
2000	**M.S.**, Geomorphology, Peking University, China
2005	**Ph.D.**, Biostatistics, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland

Employment:

2006 – present	Biostatistics Core, Masonic Cancer Center, U of M
2006 – 2013	**Assistant Professor**, non-tenure track Biostatistics, School of Public Health, U of M
2013 – 2020	**Associate Professor**, non-tenure track Biostatistics, School of Public Health, U of M
2020 – present	**Professor**, non-tenure track Biostatistics, School of Public Health, U of M

Method Research

- Recurrent event, survival data, and correlated data
- Design and analysis of clinical trials
- Statistical research motivated by collaboration experience. For example,
  - recurrent infections in BMT patients
  - correlated data in cigarette purchase task
  - AUC of biomarkers for smokers

PI'ed grants for method research:

- NIH/NIMH R03, Statistical methods for bivariate alternating recurrent event data, 2017 - 2019
- NIH/NCI R03, Statistical methods for analyzing data of recurrent infections after hematopoietic cell transplantation, 2014 - 2016
- Grant-In-Aid, Quantile regression models for recurrent gap time data, 2012 - 2014
- Minnesota Medical Foundation, Statistical methods for measuring relative reinforcing efficacy using drug purchase task instrument data, 2012 – 2013

Collaborative Research

- Tobacco and chemoprevention

Ongoing:

- NIH/NCI P01, Consortium on Methods Evaluating Tobacco (COMET): filter ventilation and product standards (Hatsukami/Shields, Core C: Luo), 2017-2023
- FDA/NIDA U54, Evaluating New Nicotine Standards for Cigarettes (Donny/Hatsukami, Core C: Koopmeiners), 2017-2023
- NIH/NCI R01, Lung cancer prevention and treatment by targeting ALDH1 and CD44 expressing putative lung cancer stem cells (PI: Kassie), 2019-2024
- NIH/NIDA R01, Impact of sugars on tobacco product toxicity and abuse liability (Stepanov/Hatsukami), 2020-2023
- NIH/NIEHS R01, Biomarker phenotypes of air pollution and cancer risk in India (Stepanov/Dikshit), 2021-2026
- NIH R01, Evaluating cigarette relighting behavior: Prevalence, correlates, toxicant exposure, and implications for cessation (Steinberg/Stepanov/Heckman), 2022-2026 (funded)

Completed (selected):

- NIH/NHLBI R01, Enhancing smoking cessation in homeless populations (Okuyemi/Pratt), 2014 - 2019
- NIH/NCMHD P60, University of Minnesota Center for Health Disparities Research, Engagement and Training (CeHDRET) (Ahluwalia, Proj 1: Thomas), 2009 - 2016
- NIH/NHLBI R01, Enhancing Quit and Win Contests to improve cessation among college smokers (Thomas), 2009 - 2014

Collaborative Research (Continued)

- BMT and Immunology

  Ongoing:

  - NIH/NCI P01 NK cells, their receptors and unrelated donor transplant (Miller, Core B: Le), 2021 - 2026
  - NIH/NCI R35, Viral priming and targeting NK cells against solid tumor malignancies (Miller), 2015-2022
  - DOD, Driving natural killer cell immunotherapy in the castration resistant prostate cancer setting with novel tri-specific killer engager molecules (Felices), 2020-2023
  - DOD, Leveraging low oxygen environments for improved natural killer cell immunotherapy (Kennedy), 2021-2023

  Pending:

  - NIH/NCI P01, Off-the-shelf immune effector cells for hematological malignancies (Wagner, Core B: Le), 2022-2027

Mentoring

- Clinical/translational researchers:
  - Dr. Antonella Borgatti, NIH K01, 2014-2019
  - Dr. Lucie Turcotte, CTSI Pre-K, 2014-2016
  - Dr. Christen Ebens, CTSI KL2, 2019-present
  - Serve as CTSI-Ed Annual Poster Session Judge: 2016, 2017, 2021
- PhD dissertations completed (or expected):
  - Chihyun Lee (2014)
  - Tianmeng Lyu (2018)
  - Anne Eaton (2020)
  - Sandra Castro-Pearson (2022 expected, advised jointly with Dr. Le)
- MS theses advised: 12 completed

Service (MCC and University)

- MCC
  - Cancer Protocol Review Committee (CPRC): 2009-2016, 2018-present
  - Data Safety and Monitoring Committee (DSMC): 2016-2018
- CTSI/U of M
  - CTSI-Ed Program Annual Poster Session Judge: 2016, 2017, 2021
  - Undergraduate Research Symposium Judge: 2018
- School of Public Health and Biostatistics Division (current)
  - Faculty Consultative Committee (FCC), At-Large Representative, 2020-2023
  - Faculty Salary Equity Review Committee (SERC), Co-Chair, 2022-2024
  - Diversity, Climate, and Inclusion (DCI) Committee, Co-Chair, 2019-present

Service (to the Profession)

- American Statistical Association (ASA) Twin Cities Chapter Representative to the Council of Chapters; 2018 - 2021.
- NIH Grant Reviews: Special Emphasis Panel, Tobacco Control Regulatory Research, Tobacco Regulatory Science A & Basic Science, Health Services Organization, Secondary Analyses of Existing Datasets of Tobacco Use and Health, Quality and Effectiveness; 2013-present.
- Journal Editorial Board: BMC Medicine (2015-2020), BMC Medical Research Methodology (2016-present)
- Journal Referee: 20+ journals
- DSMB for projects in other institutions: U of Iowa, U of Arizona

1. Background

2. Survival Analysis Methods Research
   - Recurrent event gap time methods
   - Intermittently observed time-dependent covariates
   - Composite endpoints with component-wise censoring
   - Graphical methods for survival endpoints in clinical trials
   - Collaborators

3. Statistical Methods for Behavioral Economics Data in Tobacco Research
   - Cigarette Purchase Task (CPT) Data
   - Left-censored mixed effects model
   - Two-part mixed effects model
   - Bayesian hierarchical model for meta-analysis
   - Collaborators and reference

Statistical Analysis Methods Research Overview

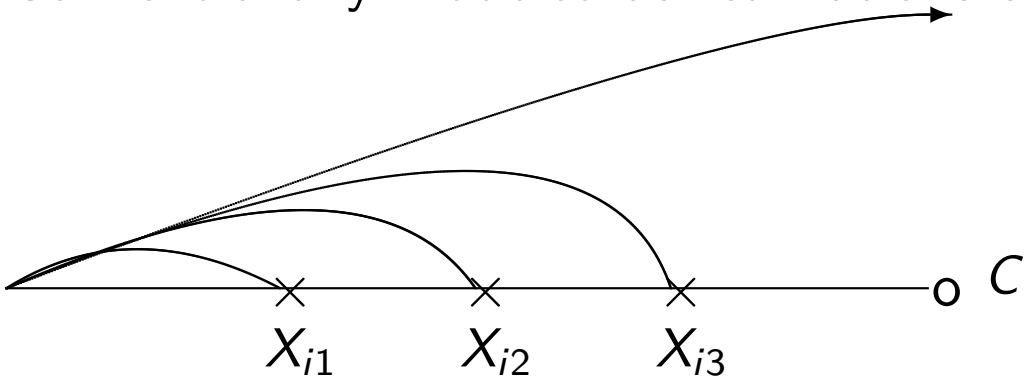
- Recurrent event gap time methods
  - Weighted risk-set & modified within-cluster resampling (i.e. "outputation")
  - Induced smoothing (for recurrent gap time AFTM)
- Intermittently observed time-dependent covariates in recurrent events models
  - Kernel smoothing
- Composite endpoints with component-wise censoring
  - Kernel smoothing (on a marker process)
- Graphical methods for survival endpoints in clinical trials
  - An "ROC" approach

Recurrent Event Gap Time Methods

Two time scales:

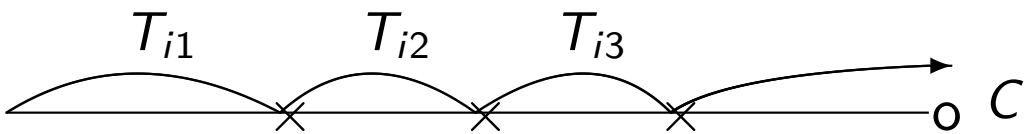
- Time since entering the study

  Conventionally modeled as realizations of counting processes.



- Time since the last event

  Focus on distribution of the gap time between consecutive events.



Recurrent Event Gap Time Methods

Tricky data structure

- Gap times are ordered and correlated
- Number of recurrent events is informative.
- Length bias
- Induced informative censoring
- Hence, clustered survival data methods (even with informative cluster size) cannot be directly applied to recurrent gap time data

Recurrent Event Gap Time Methods - WCR

Within-Cluster Resampling Method (WCR)

- In each resampling, one observation is randomly selected from each cluster to form a subsample of independent observations
- The resampling procedure is repeated a large number of times
- An estimator can be obtained through averaging over the estimates obtained from the resampled data.
- The variance of the WCR estimate can be estimated by the average of a consistent estimator of variance from each resampled dataset minus the empirical variance-covariance matrix of parameter estimates

Hoffman, Sen, and Weinberg (Biometrika, 2001), Follmann, Proschan, and Leifer (Biometrics, 2003), Williamson, Kim, Manatunga, and Addiss (SIM, 2008)

Recurrent Event Gap Time Methods - WCR

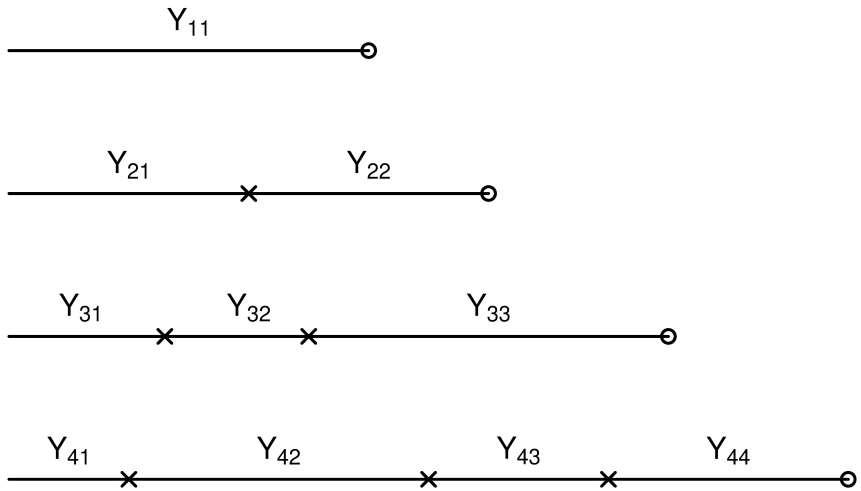
A direct application of WCR method to recurrent gap time data?

- Length bias
- Induced informative censoring

A modified WCR method for recurrent gap time data analysis under the conditions:

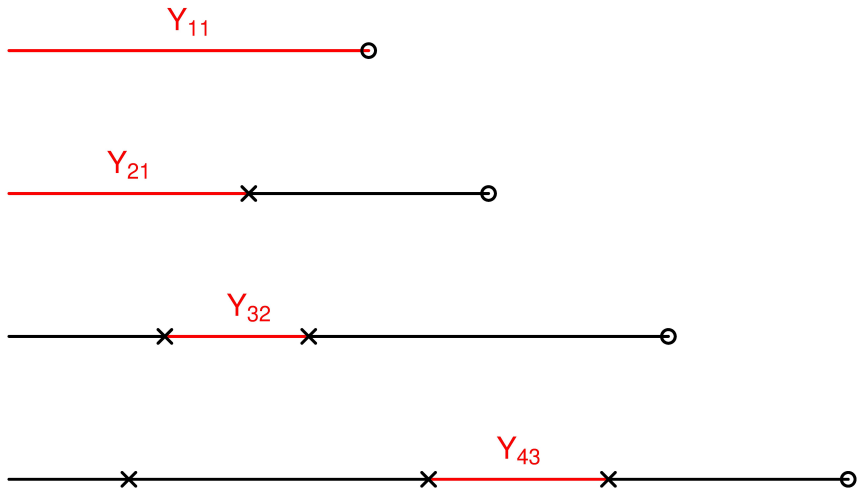
- Each individual recurrent event process is a renewal process
- The censoring time is independent of $\{T_{i1}, T_{i2}, \cdots\}$

Recurrent Event Gap Time Methods - WCR



$m_i : \sum_{j=1}^{m_i-1} T_{ij} \leq C_i$ and $\sum_{j=1}^{m_i} T_{ij} > C_i$
$m_i^* = \max\{1, m_i - 1\}$

Recurrent Event Gap Time Methods - WCR



One gap time is randomly selected from each subject to form a subsample of independent observations

Then apply Kaplan-Meier estimator, Cox model, AFT model, etc, and repeat many times.

Recurrent Event Gap Time Methods - WCR

- For each reseampling:
  - Y_{i1} is always chosen if subject i has only one censored gap time
  - Only uncensored gap times are selected if a subject has ≥ 1 events
  - The subsample consists of independent observations. Standard software can be readily applied.
- The WCR estimator can be obtained through averaging over the estimates obtained from the resampled data.

$$\hat{\beta}^{wcr} = B^{-1} \sum_{b=1}^{B} \hat{\beta}_b$$

$$\hat{\Sigma}^{wcr} = \frac{1}{B} \sum_{b=1}^{B} \hat{\Sigma}_b - \frac{1}{B-1} \sum_{b=1}^{B} (\hat{\beta}_b - \hat{\beta}^{wcr})^{\otimes 2},$$

Hoffman, Sen and Weinberg (Bmka 2001)

Recurrent Event Gap Time Methods - WRS

The Weighted Risk Set (WRS) Method

The WCR method can be used to extend univariate survival analysis methods to analyze recurrent gap time data using the idea of inverse weighting.

- Assign each observation in a risk set a weight that is proportional to the inverse of the "effective" cluster size, $1/m_i^*$
- Weighted logrank test, AFTM, model checking, etc.

Recurrent Event Gap Time Methods - Reference

Luo X, Huang C-Y. (2011). Analysis of recurrent gap time data using the weighted risk set method and the modified within-cluster resampling method. *Statistics in Medicine*, 30(4):301–311.

Huang C-Y, **Luo X**, Follmann D. (2011). A model checking method for the proportional hazards model with recurrent gap time data. *Biostatistics*, 12(3):535–547.

Luo X, Huang C-Y, Wang L. (2013). Quantile regression for recurrent gap time data. *Biometrics*, 69(2):375–385.

Lee CH, **Luo X***, Huang C-Y, DeFor T, Brunstein CG, Weisdorf DJ. (2016). Nonparametric methods for analyzing recurrent gap time data with application to infections after hematopoietic cell transplant. *Biometrics*, 72(2):535–545.

Lee CH, Huang C-Y, Xu G, **Luo X**. (2018). Semiparametric regression model for alternating recurrent event data. *Statistics in Medicine*, 37(6):996-–1008.

Lyu T, **Luo X***, Xu G, Huang C-Y. (2018). Induced smoothing for rank-based regression with recurrent gap time data. *Statistics in Medicine*, 37(7):1086–1100.

Lee CH, Huang C-Y, DeFor T, Brunstein CG, Weisdorf DJ, **Luo X**. (2019). Semiparametric regression model for recurrent bacterial infections after hematopoietic stem cell transplantation. *Statistica Sinica*, 2019 July; 29(3):1489-1509.

Intermittently Observed Time-Dependent Covariates

Now, focus on time-to-recurrent events data

Time-dependent covariates $X(t)$:

- Covariate values changing with time t.
- Usually measured at discrete time points or visits.
- Examples:
  - self-reported nicotine dependence of smokers (FTND) at monthly visits, related to the cessation outcome
  - bacterial infection status measured at regular visits, related to the risk of pharyngitis
  - Biomarkers, e.g., CD4 cell counts, related to the risk of opportunistic infections

Intermittently Observed Time-Dependent Covariates

Notation:

- $N^*(t)$: number of recurrent events occurring at or prior to time t.
- *Intensity function* of the recurrent event process $N^*(t)$:

$$\lambda(t|H(t)) = \lim_{\Delta \to 0^+} \frac{P(N^*(t+\Delta) - N^*(t) > 0 | H(t))}{\Delta} \quad (Conditional)$$

  where $H(t)$ represents the event history at or prior to time t.

- *Rate function* of the recurrent event process $N^*(t)$:

$$\lambda(t) = \lim_{\Delta \to 0^+} \frac{P(N^*(t+\Delta) - N^*(t) > 0)}{\Delta} \quad (Marginal)$$

Intermittently Observed Time-Dependent Covariates

Proportional rates model (Lin et al., 2000):

$$\lambda\{t|\mathbf{Z}_i(t)\} = \lambda_0(t)\exp\{\boldsymbol{\beta}^\top \mathbf{Z}_i(t)\}.$$

Additive rates model (Schaubel et al., 2006):

$$\lambda\{t|\mathbf{Z}_i(t)\} = \lambda_0(t) + \boldsymbol{\beta}^\top \mathbf{Z}_i(t).$$

Additive-multiplicative rates model (Liu et al., 2010):

$$\lambda\{t|\mathbf{W}_i(t)\} = \boldsymbol{\gamma}^\top \mathbf{Z}_i(t) + \exp\{\boldsymbol{\beta}^\top \mathbf{X}_i(t)\}\lambda_0(t),$$

- They all allow time-dependent covariates theoretically.
- Their model estimation procedures require the time-dependent covariates to be continuously observed throughout the entire follow up period for each subject.

Intermittently Observed Time-Dependent Covariates

Proportional rates model (Lin et al., 2000):

$$U(\beta) = n^{-1} \sum_{i=1}^{n} \int_0^{\tau} \left\{ \mathbf{Z}_i(t) - \frac{S^{(1)}(t,\beta)}{S^{(0)}(t,\beta)} \right\} dN_i(t),$$

where $S^{(k)}(t) = n^{-1} \sum_{l=1}^{n} Y_l(t) \mathbf{Z}_l(t)^{\otimes k} \exp\{\beta^{\intercal} \mathbf{Z}_l(t)\}$,
$N_i(t) = N_i^*(\min\{t, C_i\})$, $Y_i(t) = I(C_i \geq t)$.

Additive rates model (Schaubel et al., 2006):

$$U(\beta) = n^{-1} \sum_{i=1}^{n} \int_0^{\tau} \{\mathbf{Z}_i(t) - \bar{\mathbf{Z}}(t)\} dN_i(t) - \left[n^{-1} \sum_{i=1}^{n} \int_0^{\tau} \{\mathbf{Z}_i(t) - \bar{\mathbf{Z}}(t)\}^{\otimes 2} Y_i(t) dt \right] \beta,$$

where $\bar{\mathbf{Z}}(t) = \frac{n^{-1} \sum_{i=1}^{n} Y_i(t) \mathbf{Z}_i(t)}{n^{-1} \sum_{i=1}^{n} Y_i(t)}$.

Intermittently Observed Time-Dependent Covariates

Additive-multiplicative rates model (Liu et al., 2010):

$$U(\boldsymbol{\theta}) = \frac{1}{n}\sum_{i=1}^{n}\int_0^{\tau}\{\mathbf{D}_i(\boldsymbol{\theta},u) - \bar{\mathbf{D}}(\boldsymbol{\theta},u)\}\{dN_i(u) - Y_i(u)\boldsymbol{\gamma}^{\intercal}\mathbf{Z}_i(u)du\}.$$

where

$$\mathbf{D}_i(\boldsymbol{\theta},t) = \begin{pmatrix}\mathbf{Z}_i(t)/\exp\{\boldsymbol{\beta}^{\intercal}\mathbf{X}_i(t)\}\\ \mathbf{X}_i(t)\end{pmatrix},$$

$$\bar{\mathbf{D}}(\boldsymbol{\theta},t) = \frac{\frac{1}{n}\sum_{i=1}^{n} Y_i(t)\mathbf{D}_i(\boldsymbol{\theta},t)\exp\{\boldsymbol{\beta}^{\intercal}\mathbf{X}_i(t)\}}{\frac{1}{n}\sum_{i=1}^{n} Y_i(t)\exp\{\boldsymbol{\beta}^{\intercal}\mathbf{X}_i(t)\}}.$$

Intermittently Observed Time-Dependent Covariates

Additive Rates Model with Intermittently Observed Time-Dependent Covariates (Lyu et al., 2021)

By solving $U(\beta) = 0$, the estimator has a **closed-form** expression:

$$\widehat{\beta} = \left[n^{-1} \sum_{i=1}^{n} \int_0^{\tau} \{\mathbf{Z}_i(t) - \bar{\mathbf{Z}}(t)\}^{\otimes 2} Y_i(t) dt \right]^{-1} \left[n^{-1} \sum_{i=1}^{n} \int_0^{\tau} \{\mathbf{Z}_i(t) - \bar{\mathbf{Z}}(t)\} dN_i(t) \right]$$

$$= \left[\int_0^{\tau} \left(n^{-1} \sum_{i=1}^{n} Y_i(t) \frac{S^{(2)}(t)}{S^{(0)}(t)} - n^{-1} \sum_{i=1}^{n} Y_i(t) \left\{ \frac{S^{(1)}(t)}{S^{(0)}(t)} \right\}^{\otimes 2} \right) dt \right]^{-1}$$

$$\left[n^{-1} \sum_{i=1}^{n} \int_0^{\tau} \mathbf{Z}_i(t) dN_i(t) - \int_0^{\tau} \frac{S^{(1)}(t)}{S^{(0)}(t)} \left\{ n^{-1} \sum_{i=1}^{n} dN_i(t) \right\} \right].$$

Replace $\frac{S^{(k)}(t)}{S^{(0)}(t)}$ with $\frac{\widehat{S}_h^{(k)}(t)}{\widehat{S}_h^{(0)}(t)}$:

$$\frac{\widehat{S}_h^{(k)}(t)}{\widehat{S}_h^{(0)}(t)} = \frac{n^{-1} \sum_{i=1}^{n} \int_0^{\tau} K_h(t-u) Y_i(u) \mathbf{Z}_i(u)^{\otimes k} dO_i(u)}{n^{-1} \sum_{i=1}^{n} \int_0^{\tau} K_h(t-u) Y_i(u) dO_i(u)}$$

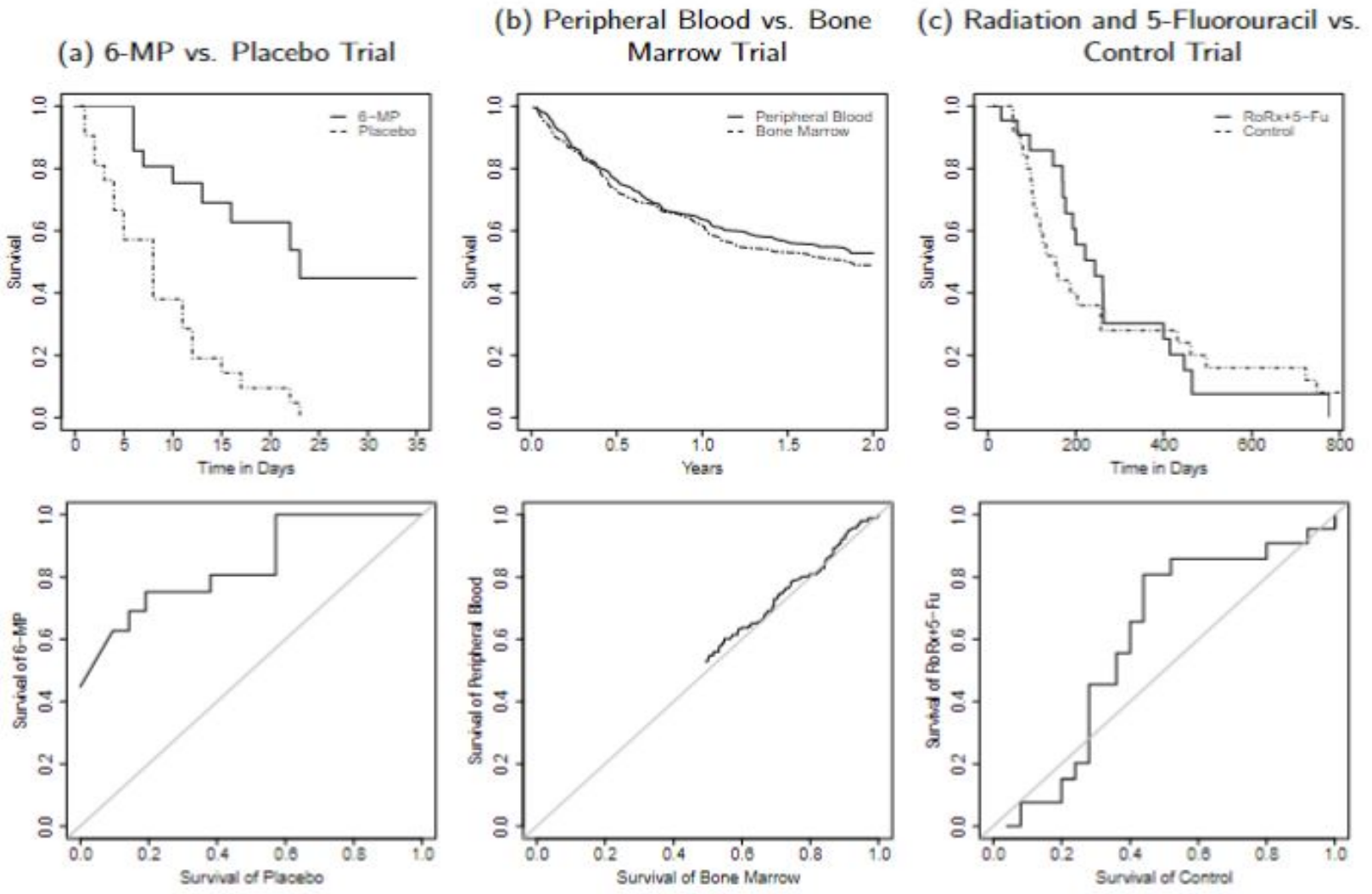
Kernel Smoothing Methods - Reference

Lyu T, **Luo X**, Huang C-Y, Sun Y. (2021). Additive rates model for recurrent event data with intermittently observed time-dependent covariates. *Statistical Methods in Medical Research*, 2021 October; 30(10):2239–2255.

Lyu T, **Luo X**, Sun Y. (2021). Additive-multiplicative rates model for recurrent event data with intermittently observed time-dependent covariates. *Journal of Data Science*, 2021 Nov 4; 19(4), 615–633.

Eaton A, Sun Y, Neaton J, **Luo X**. (2022+). Nonparametric estimation in an illness-death model with component-wise censoring. *Biometrics*, 2021 Apr 29. doi: 10.1111/biom.13482. Online ahead of print.

Graphical Methods for Survival Data in Clinical Trials



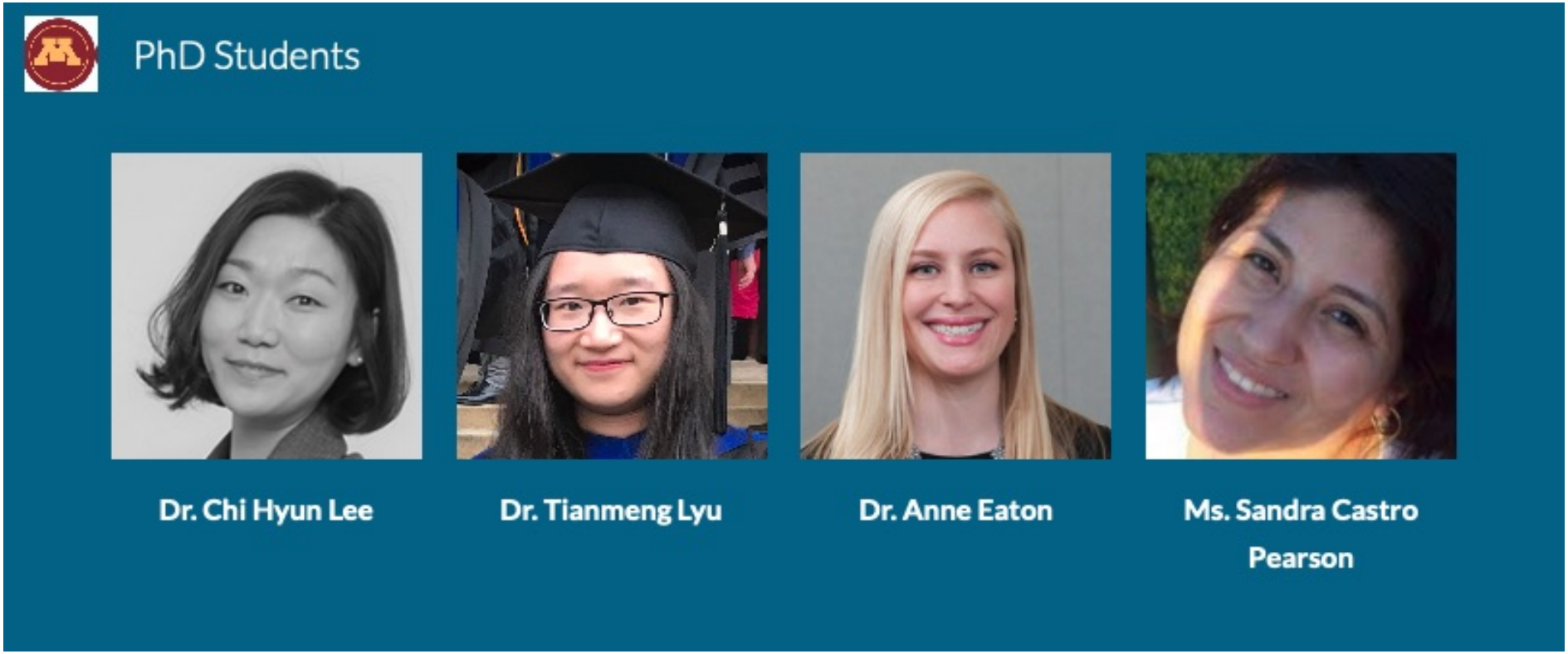
Graphical Methods for Survival Data - Reference

Castro-Pearson S, Le C, **Luo X**. (2022). Two-sample survival probability curves: A graphical approach for the analysis of time to event data in clinical trials. *Contemporary Clinical Trials*, 2022 Apr 1; 115:106707.

Castro-Pearson S, **Luo X**, Le C, et al. A graphical approach for the analysis of time to event data with competing risks in clinical trials. (Under preparation)

Castro-Pearson S, Le C, **Luo X**, et al. Monitoring randomization procedures in clinical trials. (Under preparation)

Collaborators in Survival Analysis



Collaborators

in Survival Analysis

Dr. Chiung-Yu Huang

UCSF

Dr. Lan Wang

University of Miami

Dr. Gongjun Xu

University of Michigan

Dr. Yifei Sun

Collumbia University

Dr. Chap T. Le

University of Minnesota

Statistical Methods for Behavioral Economics Data in Tobacco Research

Overview

- Family Smoking Prevention and Tobacco Control Act (2009) and tobacco regulatory science
- Product and liability and relative reinforcing efficacy (RRE) and Cigarette Purchase Task (CPT) data
- Left-censored mixed effects model
- Two-part mixed effects model
- Bayesian hierarchical model for meta-analysis

Family Smoking Prevention and Tobacco Control Act and Tobacco Regulatory Science

Family Smoking Prevention and Tobacco Control Act is a federal statute signed into law in June 22, 2009. The Act gives the Food and Drug Administration (FDA) the power to regulate the tobacco industry.

- Goal: discouraging minors and young adults from smoking
- New warnings and labels on packaging and advertisements
- Bans flavored cigarettes, places limits on the advertising of products to minors
- Requires tobacco companies to seek FDA approval for new products

Family Smoking Prevention and Tobacco Control Act and Tobacco Regulatory Science

After the Family Smoking and Prevention and Tobacco Control Act was passed, tobacco research has growing even stronger.

- A new research territory is called tobacco regulatory science
- One of the fast growing research areas is studying the Relative Reinforcing Efficacy (RRE) (a central concept in psychopharmacology) of tobacco products
- The drug purchase task (e.g., cigarette purchase task/CPT) is a frequently used instrument for measuring RRE of a product.
- As part of those efforts, my colleagues and I have focused on the modelling and data analysis of CPT data.

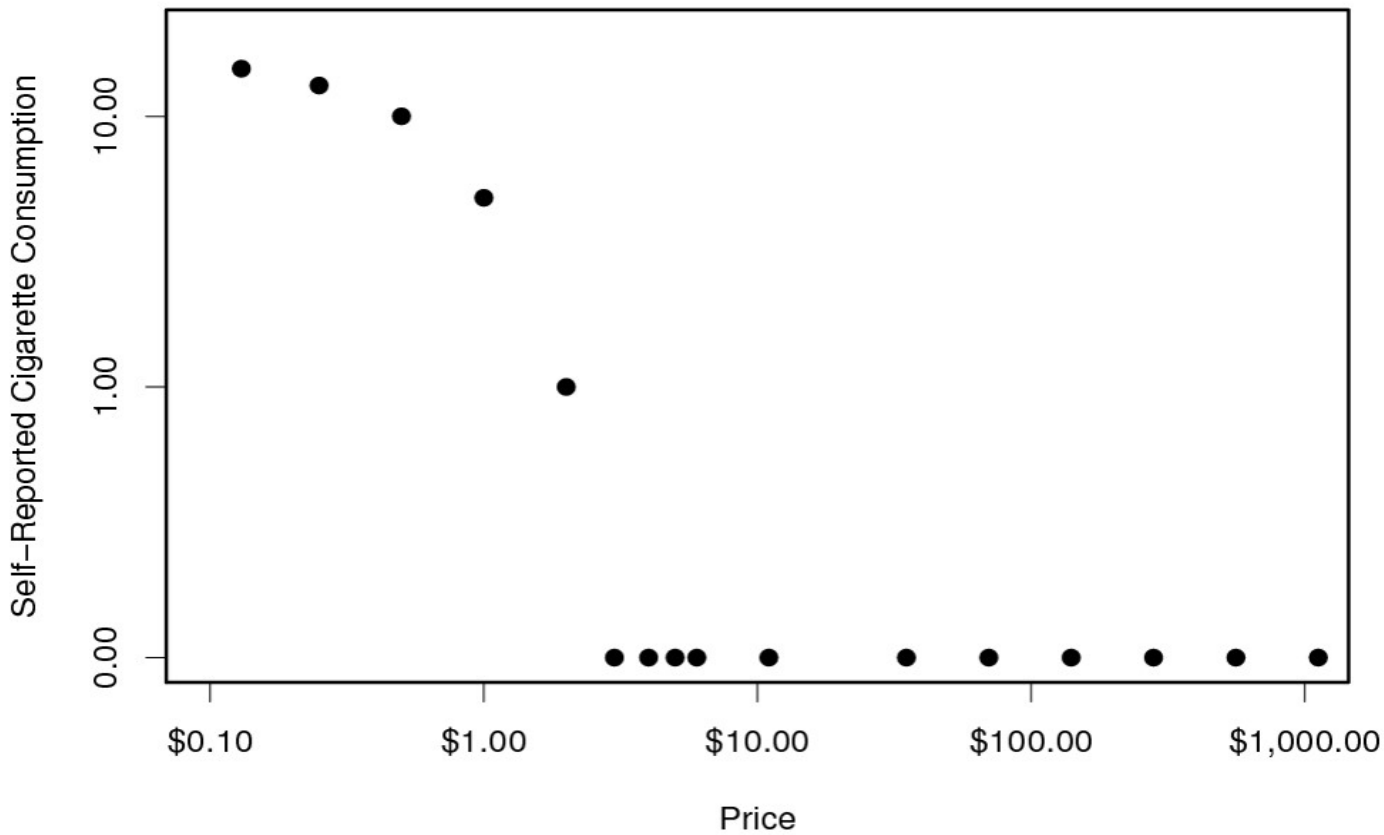
Statistical Methods for Behavioral Economics Data in Tobacco Research

Cigarette Purchase Task (CPT) Data

Imagine a TYPICAL DAY during which you smoke. The following questions ask how many cigarettes you would consume if they cost various amounts of money. Assume the following:

a. **Available cigarettes are your favorite brand**
b. **You have the same income/savings that you have now**
c. **You have NO ACCESS to any cigarettes or nicotine products other than those offered at these prices**
d. **You would consume cigarettes that you request on that day (in other words, no stockpiling)**

1. How many cigarettes would you smoke per day if they were each free?	[ ]
2. How many cigarettes would you smoke per day if they were 1¢ each?	[ ]
3. How many cigarettes would you smoke per day if they were 5¢ each?	[ ]
...	
18. How many cigarettes would you smoke per day if they were $560 each?	[ ]
19. How many cigarettes would you smoke per day if they were $1,120 each?	[ ]


Self–Reported Cigarette Consumption
0.00
1.00
10.00
$0.10
$1.00
$10.00
$100.00
$1,000.00
Price

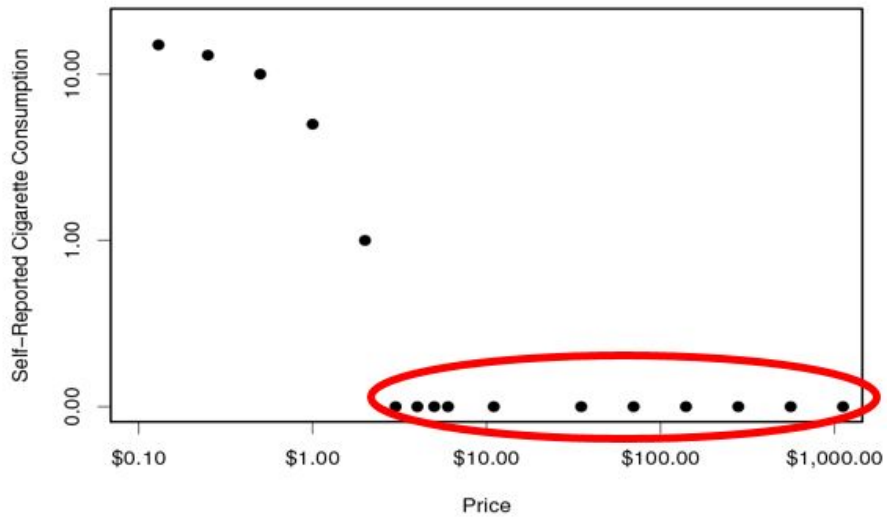
Hursh and Silberberg's exponential demand curve

$$\log Q(p) = \log Q_0 + k(e^{-\alpha p} - 1),$$

- p: price
- Q: demand for a commodity (e.g. cigarettes)
- Q_0: demand at price 0 (intensity)
- k: range of log-consumption ($= \log Q_0 - \log Q_\infty$)
- *elasticity*$(p) = \partial \log Q(p) / \partial \log p$
- *Omax* $= \max\{Q(p) \cdot p\}$: maximum expenditure
- *Pmax*: price at which *Omax* is achieved
- Breakpoint: the first price for which the consumption is zero

Hursh and Silberberg's exponential demand curve (continued)

- In application: for each subject's data $(Q_{ij}, p_j), j = 1, 2, \ldots,$ fit a nonlinear regression model to get one set of estimated parameters for each subject (overparameterized).
- How to deal with zero demand (when $p \geq$ *breakpoint*)? The log of zero is negative infinity. Ignoring zero demand or arbitrarily imputing a small value will cause bias.



Left-Censored Mixed Effects Model (Liao et al., 2013)

Goal: Fit the exponential demand curve to the data without having to delete or impute data

Assumption: the self-reported zero consumption at and beyond the breakpoint is a small nonzero consumption amount below a certain threshold (ω) that smokers do not bother to report - a limit of detection (LOD) or left censoring problem!

Model: A nonlinear mixed effects model

$$\log Q_{ij} = \mu_{ij} + \epsilon_{ij} = \log Q_{0i} + k(e^{-\alpha_i p_j} - 1) + \epsilon_{ij},$$

where

- $(\log Q_{0i}, \alpha_i)'$ is the vector of random effect (MVN with mean $(\mu_0, \mu_\alpha)'$ and variances $(\sigma_0^2, \sigma_\alpha^2)$ and covariance $\rho\sigma_0\sigma_\alpha$)
- ϵ_{ij} is the measurement error (MVN with mean 0 and variance σ_e^2).

Left-Censored Mixed Effects Model (continued)

The likelihood function for the left-censored mixed effects model has the same form as that for a correlated survival data with Type I left censoring and log-normal survival time (Klein & Moeschberger, 2003).

$$\prod_{i=1}^{N} \int \prod_{j=1}^{n_i} \left\{ \Phi\left(\frac{\log \omega - \mu_{ij}}{\sigma_e} \right) \right\}^{\delta_{ij}} \left\{ \sigma_e^{-1} \phi\left(\frac{\log Q_{ij} - \mu_{ij}}{\sigma_e} \right) \right\}^{1-\delta_{ij}} \phi(z_i) dz_i,$$

where

- $z_i = (\log Q_{0i} - \mu_0)/\sigma_0$
- δ_{ij} is the censoring indicator (=0 if $Q_{ij} \geq \omega$; =1 if $< \omega$)
- Φ and ϕ are cdf and pdf of standard normal, respectively

Two-Part Mixed Effects Model (Zhao et al., 2016)

Motivation: ω in the previous model is arbitrary; δ_{ij} can be treated as a binary outcome, with a physical meaning (cessation), and can be modeled with a logistic regression

Part I model:

$$logit(\pi_{ij}) = \beta_0 + \beta_1 f(p_j) + a_i,$$

where $\pi_{ij} = \Pr(\delta_{ij} = 1)$, and $f()$ is a proper function of price.

Part II model for $Q_{ij} > 0$ follows the same mixed effects model form as previously

$$\log Q_{ij} = \mu_{ij} + \epsilon_{ij} = (\log Q_0 + b_i) + k(e^{-(c+c_i)p_j} - 1) + \epsilon_{ij},$$

where $\gamma_i = (a_i, b_i, c_i)'$ is the vector of random effects following a MVN with mean zero.

Two-Part Mixed Effects Model (continued)

The likelihood function for two-part mixed effects model

$$\prod_{i=1}^{N} \int \prod_{j=1}^{n_i} (\pi_{ij})^{\delta_{ij}} \left\{ (1-\pi_{ij}) \sigma_e^{-1} \phi \left(\frac{\log Q_{ij} - \mu_{ij}}{\sigma_e} \right) \right\}^{1-\delta_{ij}} \phi(\gamma_i) d\gamma_i,$$

- Estimation can be carried out using the NLMIXED procedure in SAS.
- The estimated marginal parameters and the predicted random effects can be easily obtained to calculate each subject's intensity, Omax, Pmax, breakpoint, α value, etc.

Illustration with the Quit and Win Data (Thomas et al., 2016)

Quit and Win Study: 1,217 college smokers recruited from 19 two- or four-year colleges, studying two interventions (single vs. multiple contests and counseling vs. no counseling) on improving cessation
Data used: CPT data at baseline

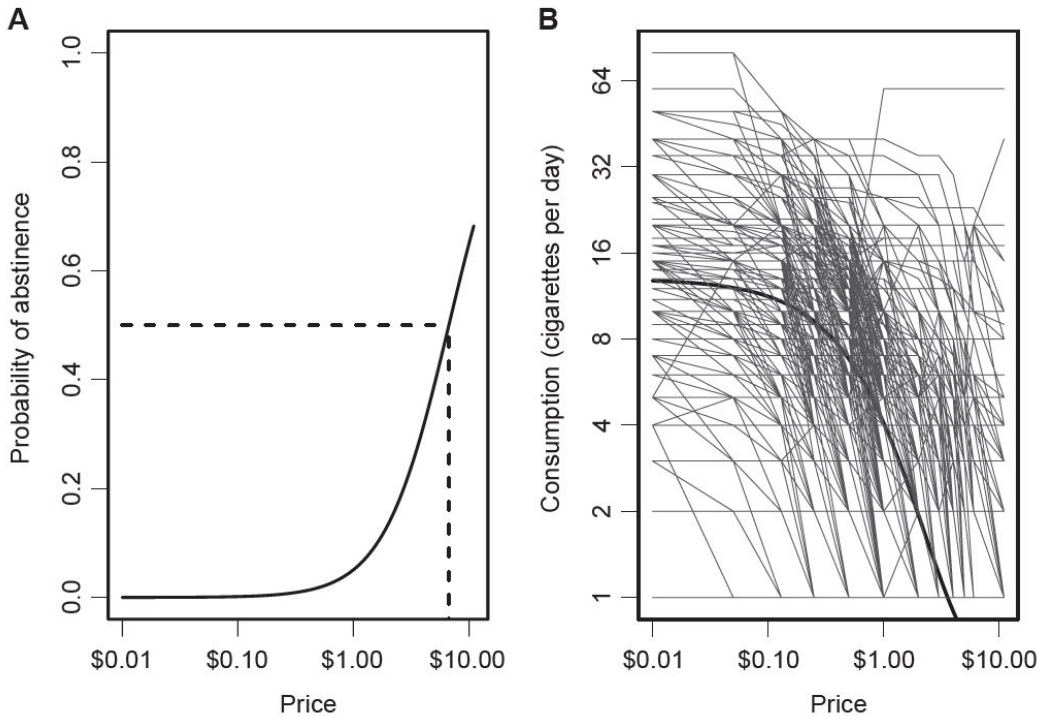


Illustration with the Quit and Win Data (continued)

Two demand parameters, intensity and breakpoint, significantly predicted reduction in cigarettes smoked per day (CPD) at 6 month among those who failed to quit.

Percent reduction in cigarettes per day among smokers at 6 months

Demand indices	β	95% CI	Type III test statistics	*p*-value
Intensity	–4.4%	(–5.4%, –3.3%)	71.40	<.0001
Omax	–0.2%	(–0.4%, 0.0%)	2.83	.09
Pmax	0.0%	(–3.7%, 3.9%)	0.00	.96
α	–2.5%	(–34%, 29%)	0.02	.88
Breakpoint	–4.1%	(–5.6%, –2.6%)	29.17	<.0001

Bayesian Hierarchical Model for Meta-Analysis (Zhang et al., in press)

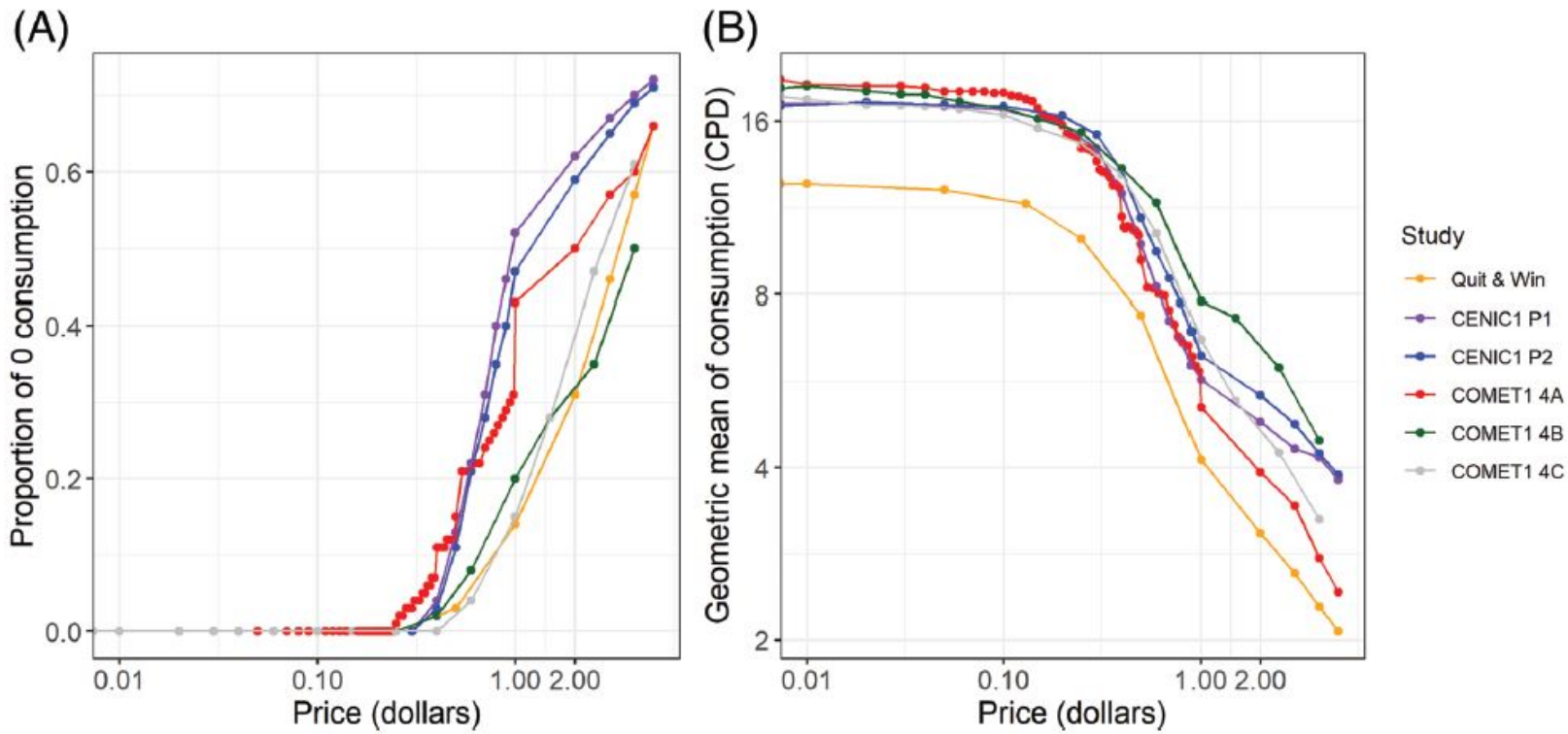
Six recent, multi-center tobacco studies coordinated by the University of Minnesota and designed and analyzed by the MCC Biostatistics Core.

Study	Full sample size	Analysis Sample size[a,b]	Age ≤ 25[c]
Quit & Win	1217	1214 (> 99%)	700 (58%)
CENIC1 P1	840	839 (> 99%)	137 (16%)
CENIC1 P2	1250	1227 (98%)	102 (8%)
COMET1 4A	224	179 (80%)	22 (12%)
COMET1 4B	211	181 (86%)	15 (8%)
COMET1 4C	295	223 (76%)	24 (11%)

Motivation: Extend the two-part mixed effects model by Zhao et al. by using a Bayesian hierarchical model to estimate the study-specific and population-averaged parameters simultaneously.

Bayesian Hierarchical Model for Meta-Analysis (continued)

Summary of the observed data of each study

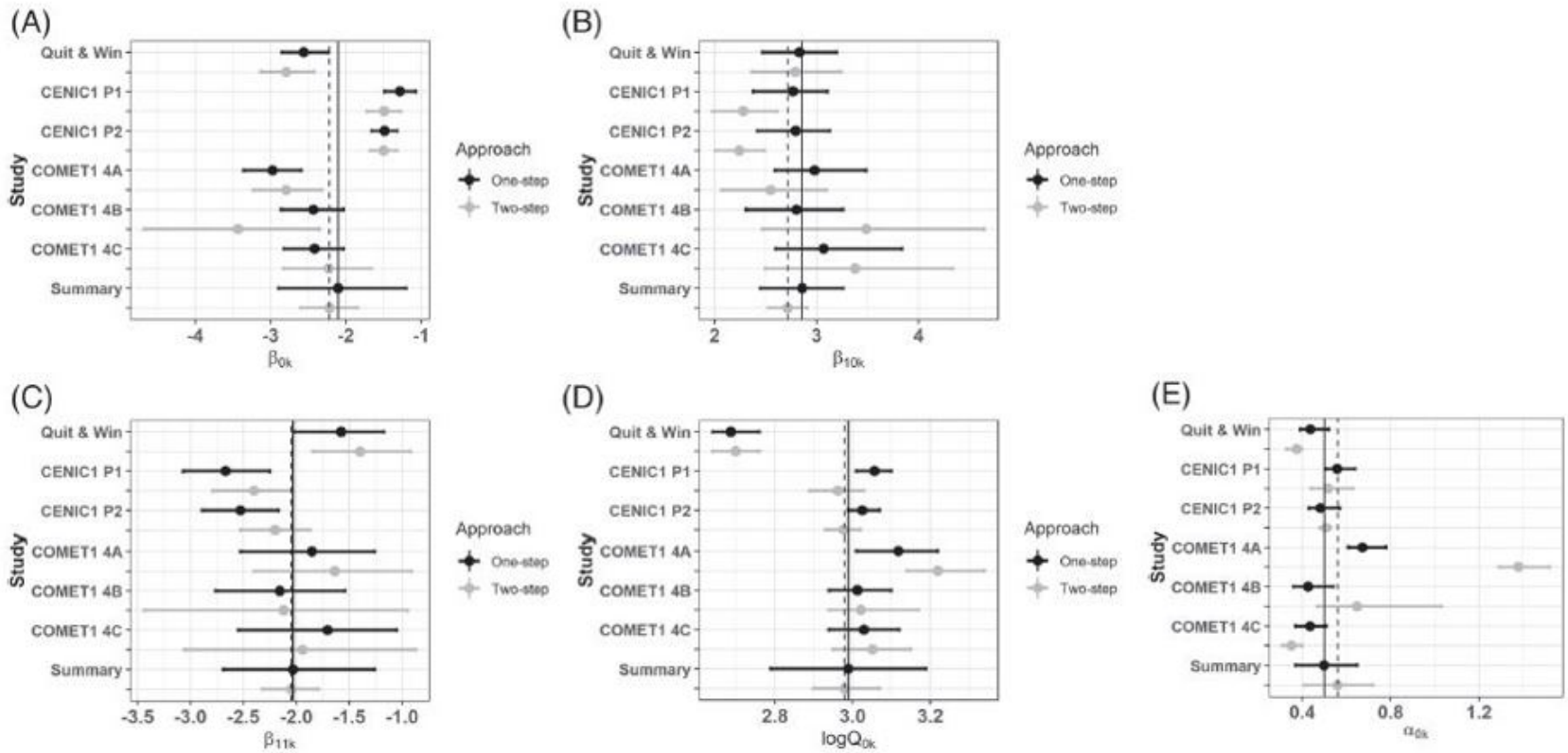


Bayesian Hierarchical Model for Meta-Analysis (continued)

- Compared with the Zhao et al. model, the within-study model of the proposed Bayesian hierarchical model adopts a more flexible shape (splines) for the function of price, $f(p)$ in the Part I model.
- Subject-level covariates are incorporated in the within-study model.
- Between-study model allows heterogeneity across studies, by assuming study-specific parameters to follow normal distributions.
- Noninformative priors were used.
- The likelihood and posterior distribution can be derived.
- MCMC was used for posterior computation, carried out using JAGS and the `rjags` package in R.

Bayesian Hierarchical Model for Meta-Analysis (continued)

Compared with the two-step approach



- Shrinks toward the population-level parameters

Bayesian Hierarchical Model for Meta-Analysis (continued)

- This work is meant to be a proof of concept for demonstrating the utility of the Bayesian framework for the analysis of complex data
- The proposed model for CPT data contains a substantial number of parameters, may require the study of its sensitivity to choices of priors
- The RRE indicies which are not directly from the model parameters, e.g. Omax and Pmax, can be easily obtained with posterior samples

Collaborators in CPT Methods

- Biostatistics Students
  - Dr. Wenjie Liao
  - Ms. Tingting Zhao
  - Mr. Shengwei Zhang
- Biostatistics Faculty
  - Dr. Chap T. Le
  - Dr. Haitao Chu
  - Dr. Yen-Yi Ho (South Carolina)
- U of M, Medical School and Masonic Cancer Center
  - Dr. Janet L. Thomas
  - Dr. Dorothy K. Hatsukami
- Other Universities
  - Dr. Leonard H. Epstein (University at Buffalo)
  - Dr. Eric C. Donny (Wake Forest)
  - Dr. Warren K. Bickel (Virginia Tech)
  - Dr. Tracy T. Smith (University of South Carolina)

Reference (CPT Methods)

Liao W, **Luo X***, Le C, Chu H, Epstein LH, Yu J, Ahluwalis JS, Thomas J. (2013). Analysis of cigarette purchase task instrument data with a left-censored mixed effects model. *Experimental and Clinical Psychopharmacology*, 21(2):124–132.

Zhao T, **Luo X***, Chu H, Le CT, Epstein LH, Thomas JL. (2016). A two-part mixed effects model for cigarette purchase task data. *Journal of the Experimental Analysis of Behavior*, 106(3):242–253.

Ho YY, Vo TN, Chu H, LeSage M, **Luo X**, Virnig B, Le CT. (2018). A Bayesian hierarchical model for demand curve analysis. *Statistical Methods in Medical Research*, 27(7):2038–2049.

Zhang S, Chu H, Bickel WK, Le CT, Smith TT, Thomas JL, Donny EC, Hatsukami DK, **Luo X***. (2022+). A Bayesian hierarchical model for individual participant data meta-analysis of demand curves. *Statistics in Medicine*, 2022 Feb 22. https://doi.org/10.1002/sim.9354. Online ahead of print.

Acknowledgement

- MCC Biostatistics Core:
  Bruce Lindgren, Todd DeFor (Metronic), Qing Cao, Ryan Shanley, Katelyn Tessier, Nathan Rubin
- Mentors in Biostat Division:
  (Official) Drs. James Hodges and Chap Le; (Unofficial) Drs. Wei Pan, James Neaton, Haitao Chu, John Connett (Prof Emeritus)
- MCC Collaborators:
  (Biostat Core) Drs. Rachel Isaksson, Saonli Basu, Joseph Koopmeiners, Lin Zhang, Ashley Petersen, Thomas Murray, Anne Eaton;
  (Tobacco and Chemoprevention) Drs. Janet Thomas, Dorothy Hatsukami, Stephen Hecht, Jasjit Ahluwalia (Brown), Larry An (Michigan), Kolawole Okuyemi (Utah), Rebekah Pratt, Fekadu Kassie, Chengguo Xing (Florida), Irina Stepanov
  (BMT and Immunology) Drs. Jeffrey Miller, John Wagner, Daniel Weisdorf, Frank Cichocki, Martin Felices, Philippa Kennedy

Thank you!