Longitudinal Analysis—Better than Ezra:
Using Nehemiah as his own Control

Karen Bandeen-Roche
Professor of Biostatistics & Medicine
Johns Hopkins University

University of Michigan OAIC RCDC Retreat
June 1, 2006

Why longitudinal data analysis (LDA)?

- Top ten reasons

10. Because it will make me look so cool

9. Because a grant reviewer will call my application “unsophisticated” if not

(I’m only creative enough to come up with two of these….)

Why LDA?

- Top four reasons

4. To inform public policy
   - Changes in disability prevalence over time
3. To study natural histories
   - Functional trajectories and their etiologies
2. To make prognoses, incorporating history
   - Cognitive status transitions
1. To progress from “association” toward “cause”
   - Intervention A or risk adoption B changes outcomes

What I Hope You'll Get Out of This

- The basic longitudinal modeling methods
- How one implements those methods
  - Key models
  - Software
- Heads up on the primary challenges
- Heads up on causality considerations

An Example
Emotional vitality and mobility

- Study: Women's Health & Aging *(n=1002; Guralnik et al., 1995)*
- Question: Does emotional vitality affect mobility trajectory?
  - Emotional vitality (X: 1 if vital; 0 ow)
    - High mastery, being happy, few depressive/anxious symptoms

Penninx et al., 2000

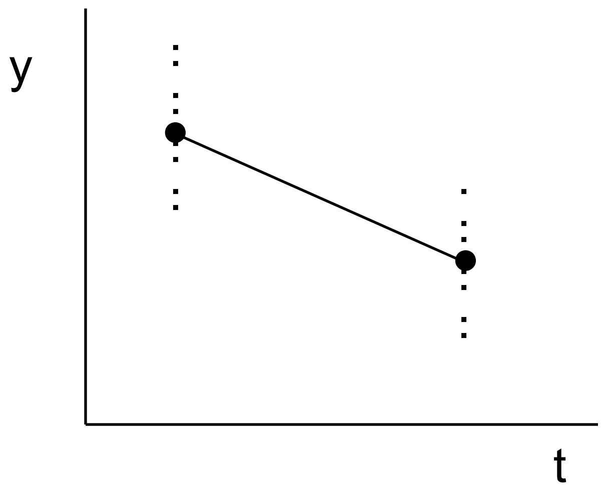
  - Mobility (Y)
    - Usual walking speed (max 2 trials)
    - Indicator of severe walking difficulty (1 if yes; 0 ow)
  - Time (T)
    - Study rounds 0-6

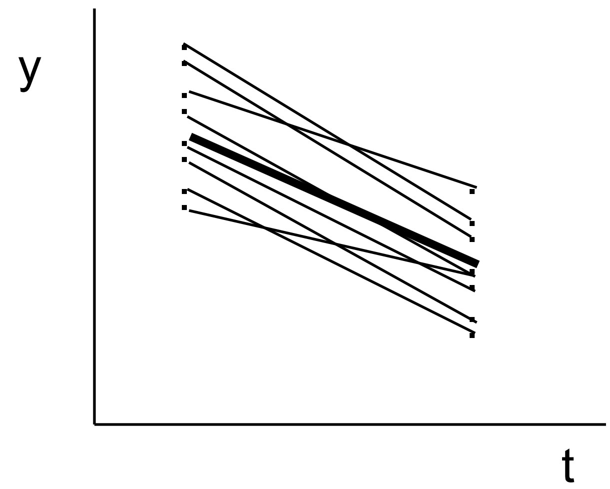
The basic longitudinal methods
Diggle, Heagerty, Liang & Zeger, 2001

- Top four reasons

4. To inform public policy
   – Population average (marginal models; GEE)
3. To study natural histories
   – Subject-specific (random effects; growth curves)
2. To make prognoses, incorporating history
   – Transitions (autoregressive & Markov models)
1. To progress from “association” toward “cause”
   – Time-varying covariates (with complexities)

Population average v. Subject-Specific





- PA: Compare populations over time
  - (Fixed) time effect = slope of the averages
- SS: Compare women **to selves** over time
  - (Fixed) time effect = average of the slopes
- Subtle point: These are equal
  - with continuous outcomes Y (linear regression); NOT otherwise
  - provided that within-person correlation is explicitly accounted for

Population-average models

- Keywords
  - Marginal models
  - GEE (Generalized Estimating Equations)

    Liang & Zeger, 1986
  - Panel analysis
- Sound bites
  - Focus usually on averages (their trajectories)
  - Serial correlation often a “nuisance”
  - “Robust”

Population-average models
Description of average trajectories

- Model—time-invariant covariates:

$$Y_{i1} = \beta_0 + \beta_1 x_i + \beta_2 t_{i1} + \beta_3 x_i \cdot t_{i1} + e_{i1}$$

$$Y_{ij} = \beta_0 + \beta_1 x_i + \beta_2 t_{ij} + \beta_3 x_i \cdot t_{ij} + e_{ij}$$

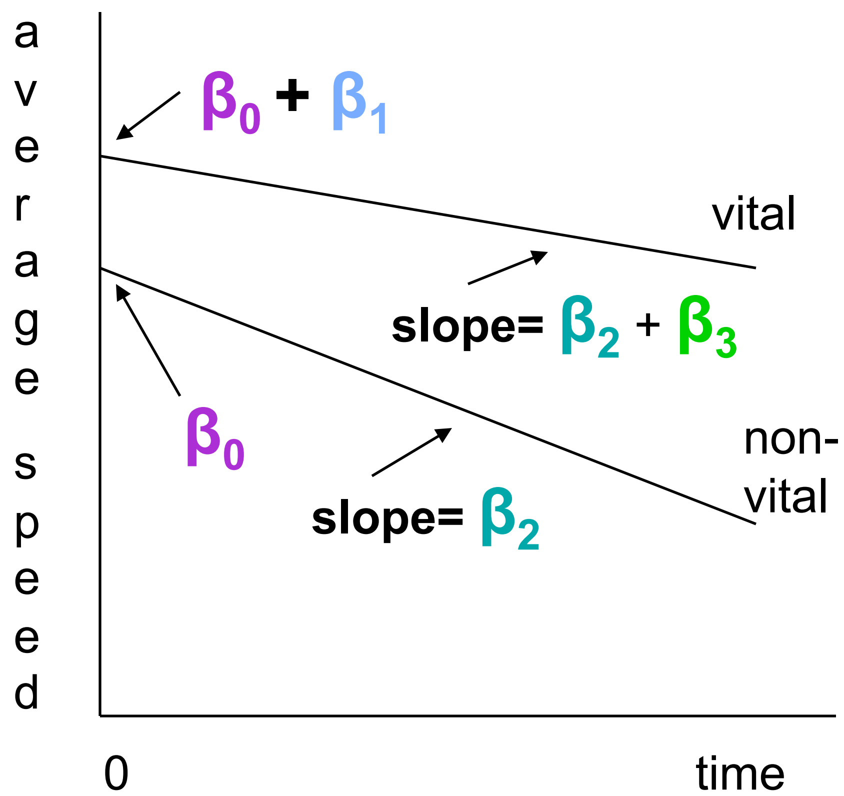
$$Y_{i7} = \beta_0 + \beta_1 x_i + \boxed{\beta_2} t_{i7} + \boxed{\beta_3} x_i \cdot t_{i7} + e_{i7}$$

rate of change in average walk speed of non-vital persons

amount rate of change in average walk speed differs between vital & non-vital persons

- Key points
  - Greek = "fixed"; Roman = variable
  - "ANCOVA" model
    - **Coding: main effects for "treatment," time; interaction**
    - Note contrast viz "change scores": more powerful

Population-average models
Pictures



- Data displays
  - Side-by-side box plots (by time, "treatment")
  - Connect-the-means plots (over time, by treatment)
  - Y versus t smoothed scatterplot, per x

$$Y_{ij} = \beta_0 + \beta_1 x_i + \beta_2 t_{ij} + \beta_3 x_i \cdot t_{ij} + e_{ij}$$

Population-average models

Treatment of serial correlation

$$Y_{i1} = \beta_0 + \beta_1 x_i + \beta_2 t_{i1} + \beta_3 x_i \cdot t_{i1} + e_{i1}$$

$$Y_{ij} = \beta_0 + \beta_1 x_i + \beta_2 t_{ij} + \beta_3 x_i \cdot t_{ij} + e_{ij}$$

$$Y_{i7} = \beta_0 + \beta_1 x_i + \beta_2 t_{i7} + \beta_3 x_i \cdot t_{i7} + \boxed{e_{i7}}$$

error: amount that speed of woman "i" differs from population average at time 7

- Key points
  - Errors are correlated within persons
  - Most software: you choose the correlation "structure"
    - "Exchangeable" – all measures equally strongly correlated
    - "Autoregressive," "banded" – measures closer in time more strongly correlated
    - "Unstructured" – as it sounds (here: 7 choose 2 = 21 ρs)
    - "Independence" – all correlations assumed = 0

Population-average models Categorical outcomes

$$F(\text{mean}[Y_{i1}]) = \beta_0 + \beta_1 x_i + \beta_2 t_{i1} + \beta_3 x_i \cdot t_{i1} \cancel{+ e_{i1}}$$

$$F(\text{mean}[Y_{ij}]) = \beta_0 + \beta_1 x_i + \beta_2 t_{ij} + \beta_3 x_i \cdot t_{ij} \cancel{+ e_{ij}}$$

$$F(\text{mean}[Y_{i7}]) = \beta_0 + \beta_1 x_i + \beta_2 t_{i7} + \beta_3 x_i \cdot t_{i7} \cancel{+ e_{i7}}$$

- F(mean[Y_{ij}]) not so complicated
  - e.g. binary Y: (mean[Y_{ij}]) = Prob{Y_{ij} = 1}, say p_{ij}
    - F = logit(p) = log(p/[1-p]) = log(odds)
    - Longitudinal logistic regression
  - Most software: you choose F ("link")
    - Logit, log, probit, complementary log-log, identity
  - Population interpretation, correlation choices: same

Population-average models: Fitting

- Software
  - SAS: GENMOD (GEE); MIXED, repeated (MLE)
  - SPSS: Advanced model package
  - Stata: xtgee (GEE); xtreg (MLE)
- GEE versus MLE (maximum likelihood est.)
  - Both: accurate coefficient estimates whether or not correlation structure choice is correct
  - GEE: standard errors also accurate, regardless
  - MLE: More powerful if choice is correct

Subject-specific models

- Keywords
  - Mixed effects, growth curves, multi-level
  - Mixed model; hierarchical (linear) model GEE

    Laird & Ware, 1982; Raudenbush & Bryk, 1986
  - Random coefficient model
- Sound bites
  - Focus usually on individual trajectories
  - “Heterogeneity”: variability of trajectories
  - Assumptions are made, and may matter

Subject-specific models
Average & individual trajectories

- Model—time-invariant covariates:

$$Y_{i1} = \beta_0 + b_{0i} + \beta_1 x_i + \beta_2 t_{i1} + b_{2i}\, t_{i1} + \beta_3 x_i \cdot t_{i1} + e_{i1}$$

$$Y_{ij} = \beta_0 + b_{0i} + \beta_1 x_i + \dot{\beta}_2 t_{ij} + b_{2i}\, t_{ij} + \beta_3 x_i \cdot t_{ij} + e_{ij}$$

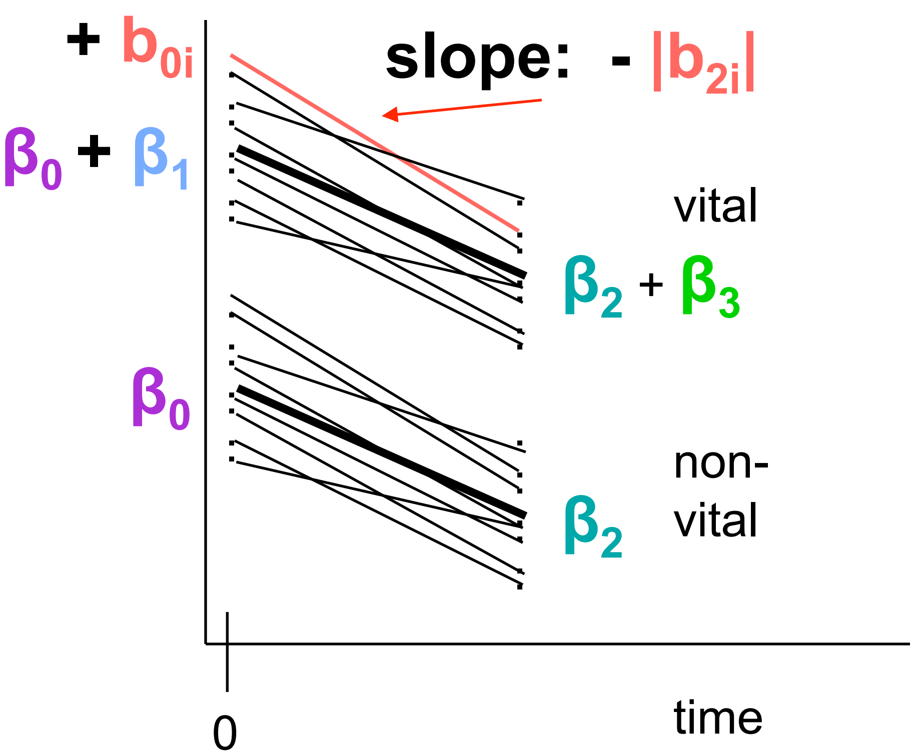
$$Y_{i7} = \beta_0 + \boxed{b_{0i}} + \beta_1 x_i + \dot{\beta}_2 t_{i7} + \boxed{b_{2i}}\, t_{ij} + \beta_3 x_i \cdot t_{i7} + e_{i7}$$

amount baseline speed for person i exceeds or falls short of the average

amount speed trajectory for person i differs from average

- Key points:
  - The additional coefficients are random
  - Modeling assumes a distribution: usually normal
    - Distribution variance characterizes "heterogeneity"
    - Heterogeneity results in within-person correlation
  - One may define correlation structure for e_{ij}s too

Subject-specific models
Pictures



- b_{0i} = random intercept
  b_{2i} = random slope
  (could define more)
- heterogeneity ⬌ spread in intercepts, slopes
- Sentinel data display: spaghetti plot
  (Ferrucci et al., 1996)

$$Y_{ij} = \beta_0 + b_{0i} + \beta_1 x_i + \beta_2 t_{ij} + b_{2i} t_{ij} + \beta_3 x_i \cdot t_{ij} + e_{ij}$$

Subject-specific models
Categorical outcomes

$$F(\mathrm{mean}[Y_{i1}|b_{0i}, b_{2i}]) = \beta_0 + b_{0i} + \beta_1 x_i + \beta_2 t_{i1} + b_{2i} t_{i1} + \beta_3 x_i \cdot t_{i1} \cancel{+ e_{i1}}$$

$$F(\mathrm{mean}[Y_{ij}|b_{0i}, b_{2i}]) \doteq \beta_0 + b_{0i} + \beta_1 x_i + \beta_2 t_{ij} + b_{2i} t_{ij} + \beta_3 x_i \cdot t_{ij} \cancel{+ e_{ij}}$$

$$F(\mathrm{mean}[Y_{i7}|b_{0i}, b_{2i}]) \doteq \beta_0 + b_{0i} + \beta_1 x_i + \beta_2 t_{i7} + b_{2i} t_{i7} + \beta_3 x_i \cdot t_{i7} \cancel{+ e_{i7}}$$

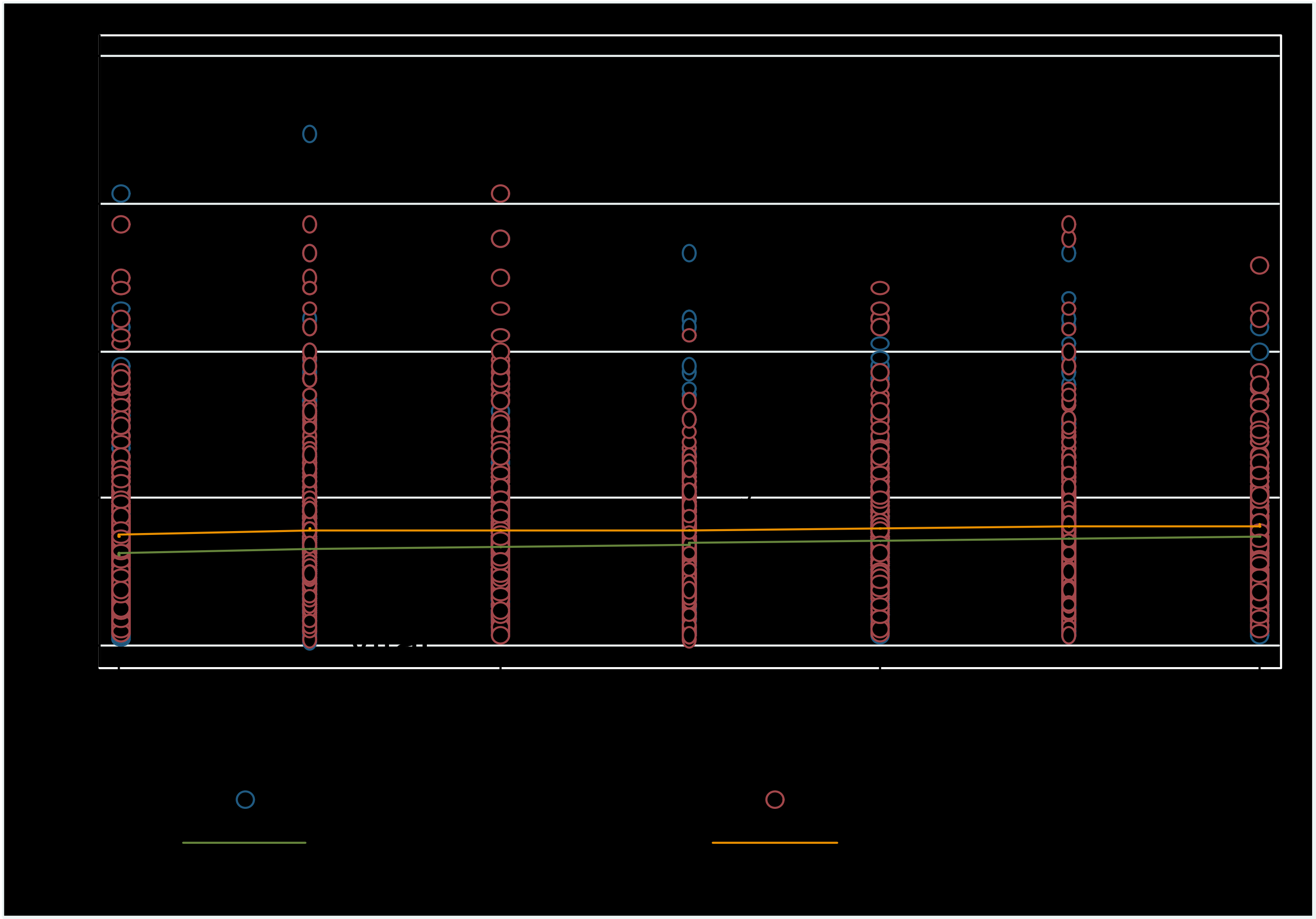
- Some fixed effect interpretations are subject specific
  - Logistic regression example: β_1 = log odds ratio comparing woman i baseline risk if vital vs non-vital
  - Only informed by data if vitality status varies with time
  - These effect estimates sensitive to distribution choice
- MLE approximation: Computationally intensive

Subject-specific models: Fitting

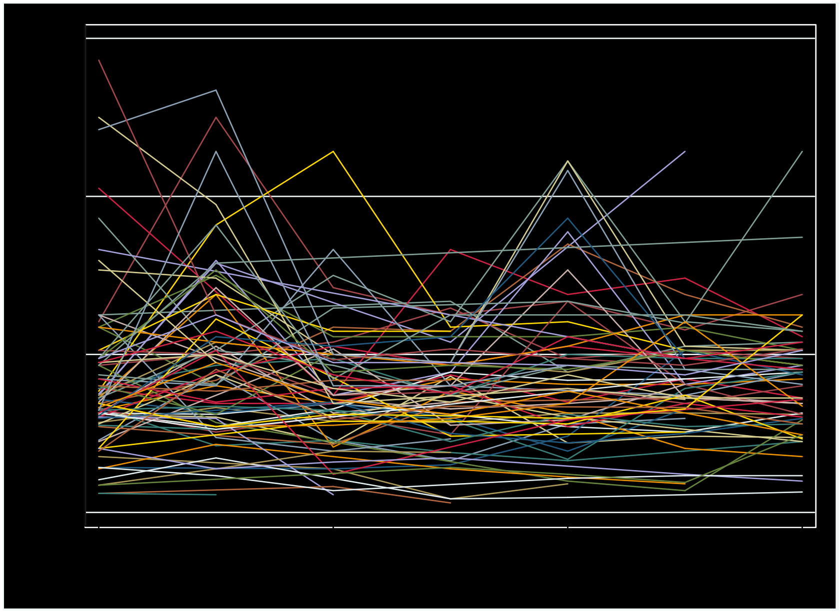
- Software
  - SAS: MIXED, random; GLIMMIX (macro);
    NLMIXED
  - SPSS: Advanced model package
  - Stata: xt… sequence
  - Other: HLM, MLWIN, Splus, R, winbugs
- Sister formulation: latent growth curve

Interlud
e

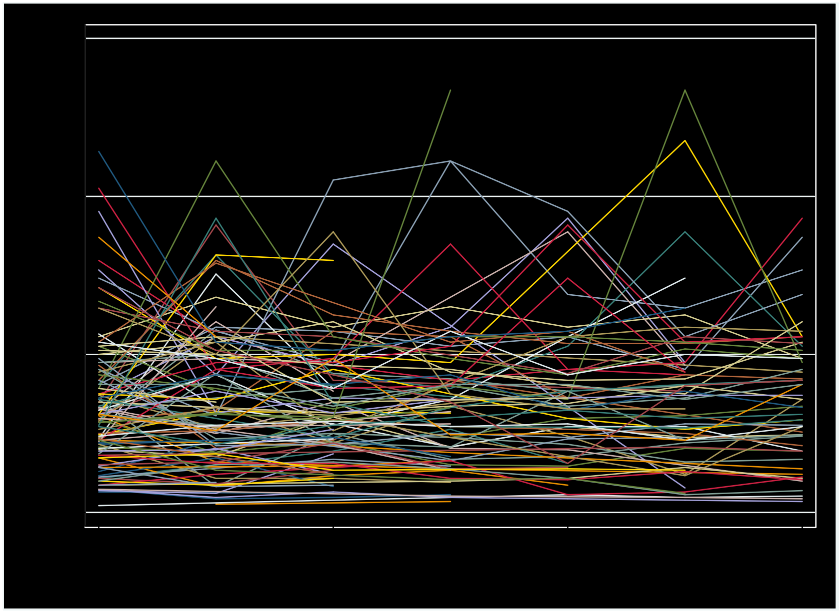
Usual Walking Speed in WHAS Panel Plot



Usual Walking Speed in WHAS Spaghetti Plots



Emotionally vital



Emotionally non-vital

Usual Walking Speed in WHAS Does vitality affect walking speed?

Parameter	ML: Independent	GEE: unstructured	ML: unstructured	ML: Random b_0 & b_1
Intercept	.58 (.010)	.63 (.035)	.57 (.012)	.58 (.012)
Vitality	**.10 (.017)**	***.075 (.050)***	**.10 (.020)**	**.10 (.020)**
Time	.0026 (.003)	**-.031 (.012)**	**-.012 (.0022)**	**-.012 (.002)**
Vit*time	-.0015 (.005)	.017 (.018)	**.0068 (.0035)**	***.0062 (.0034)***
Main effects model: Intercept, vitality results very similar to above				
Time	.0020 (.0024)	**-.0058 (.002)**	**-.0091 (.002)**	**-.0094 (.002)**

wrong

Usual Walking Speed in WHAS Heterogeneity

- Residual SD, variance: 0.167, .0280
  - Represents variability of a woman's speeds "about" her own regression line
- Intercept SD, variance: 0.276, .0762
  - "Test-retest" estimate = .076/(.076+.028)=.73
- Slope SD, variance: 0.031, .00094
  - 95% of slopes estimated within +/-.06 of ~-.01
- Intercept, slope covariance: .0020
  - Correlation=.23: better trajectories for better starters
- Unstructured correlations: .6 - >.99
  - Highest late in the study

Vitality & Walking Speed in WHAS
Summary

- Beneficial association with emotional vitality
  - Begin better by ~.1; 95% CI ~ [.06,.14]
  - Moderate evidence: Decline rate ~ halved
- Remarkable stability evidenced
  - Modest average decline
  - Heterogeneity: moderate ↓ to modest ↑
  - Stability increased with duration in study
- To advance toward “causation”: much needed
  - Control for confounders
  - Change on change

Population average v. Subject-Specific
How to choose?

- Science
- Advantages of subject-specific models
  - Characterization of heterogeneity–estimates
  - May well embody mechanisms
- Advantages of marginal models
  - More robust
    - Standard errors valid if correlation model wrong (GEE)
    - Fixed effect estimates distribution-insensitive
  - Computationally faster, more transportable (GEE)
- An MLE advantage: Missing data treatment

Why LDA?

- Top four reasons

4. To inform public policy
   – Changes in disability prevalence over time
3. To study natural histories
   – Functional trajectories and their etiologies
2. To make prognoses, incorporating history
   – Cognitive status transitions
1. To progress from “association” toward “cause”
   – Intervention A or risk adoption B changes outcomes

Some LDA & causality punch lines

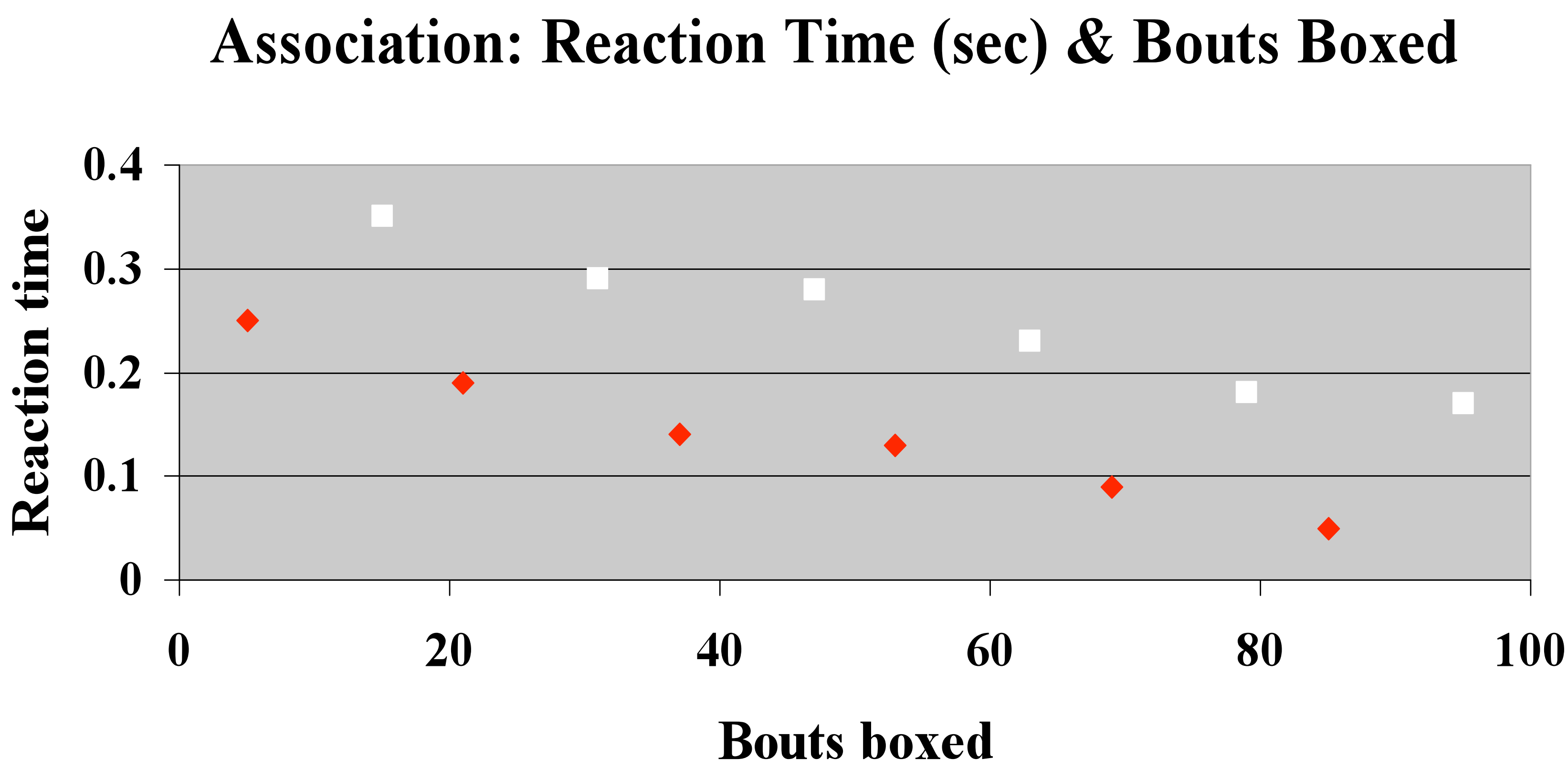
- That’s “progress from ‘association’ toward ‘cause’”
  - Temporality = one necessary component of causality
  - The others: association, isolation

    von Suppes, 1970; Bollen, 1989; Rubin, 1974

- Not all LDAs are created equal
  - Top of the hierarchy: Change-on-change
    - Change in response (Y) versus change in predictor (X)
    - Approximates “potential outcomes” observation (e.g. crossover)
  - Key = use of individuals as their own controls

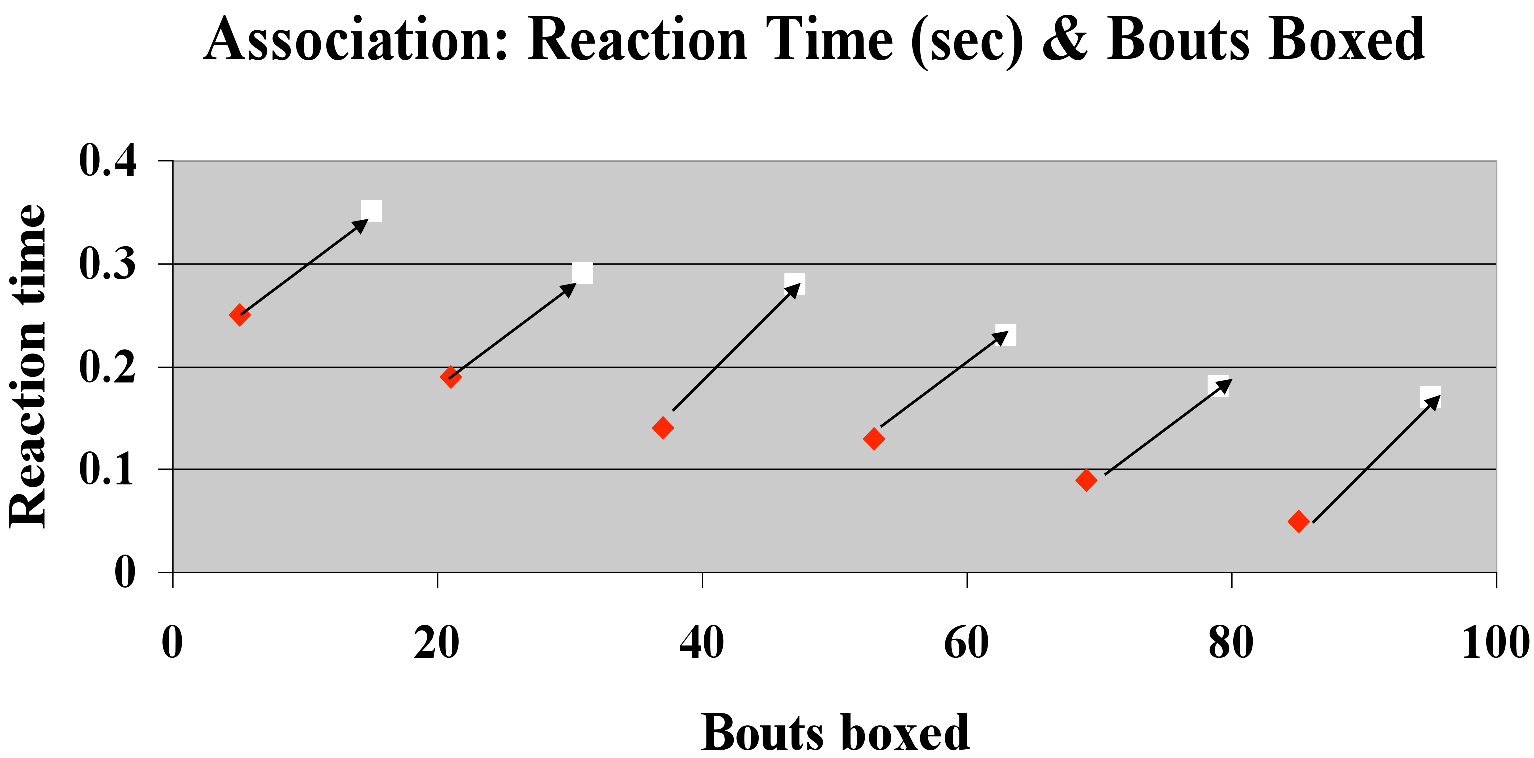
Value of change-on-change

Neuropsychological effects of amateur boxing



Value of change-on-change

Neuropsychological effects of amateur boxing



"Unlinking" model: Bandeen-Roche et al., 1999

LDA Challenge # 1
Feedback, endogeneity

- Decline in speed may erode emotional vitality… or, the vital may try harder at the measured walk test
- An issue with time-varying or invariant xs
- Solution # 1: Sophisticated modeling
  - Cross-lag, Structural, Marginal Structural
  - *Geweke, 1982; Bollen, 1989; Robins, 1986*
- Solution # 2: Transition modeling

Why LDA?

- Top four reasons

4. To inform public policy
   – Changes in disability prevalence over time
3. To study natural histories
   – Functional trajectories and their etiologies
2. To make prognoses, incorporating history
   – Cognitive status transitions
1. To progress from "association" toward "cause"
   – Intervention A or risk adoption B changes outcomes

Transition Models

- Basic idea: control model for current outcome on all past outcomes
  - Autoregressive errors
  - Modify marginal model to include past "Y"s as predictors in model for Y_{it}

- Often assumed: current outcome only depends on the one most immediately past
  - Model for Y_{it} includes Y_{it-1} but no other Ys
  - "First order Markov"

Beckett et al.,1996

LDA Challenge # 2
Dropout, Missing Data

- <u>The issue</u>: Those "missing" may differ systematically from those observed
  - Sicker?
  - Less emotionally vital?
  - Functionally declining?

- Findings' accuracy, precision may suffer

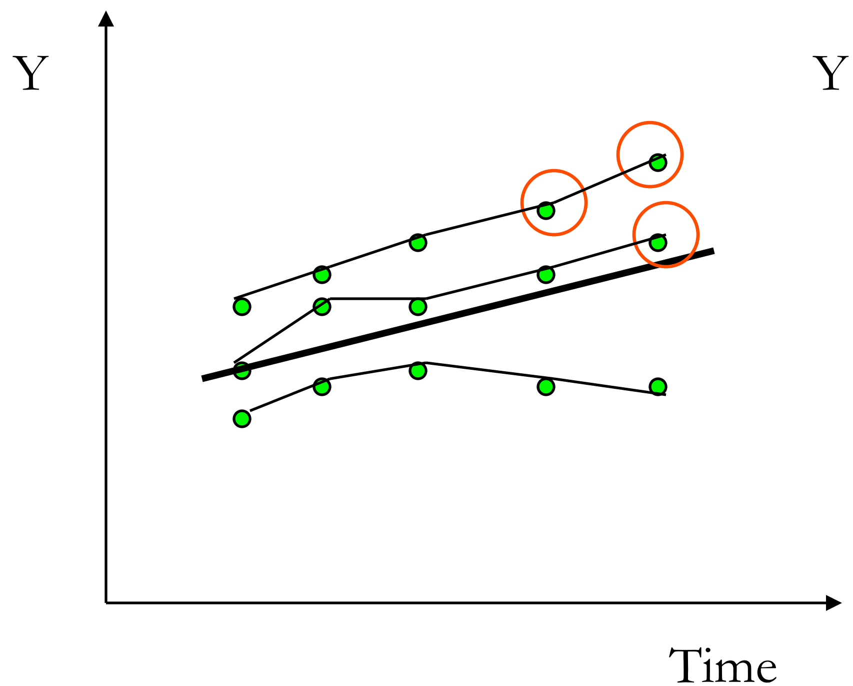
Missing data, and Missing data

Rubin, 1976; Little & Rubin 1989

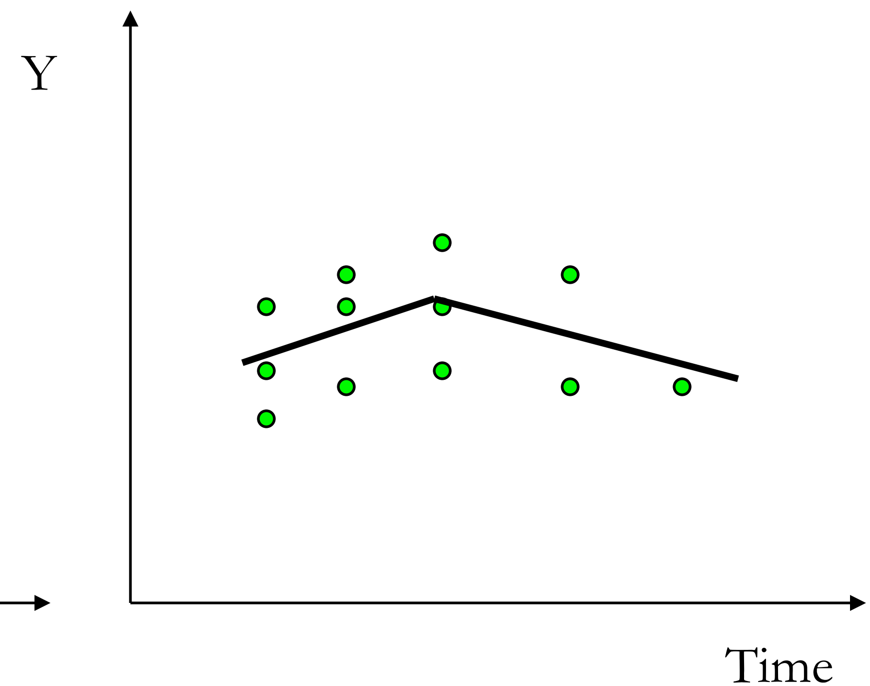
- A standard hierarchy:
  - Missing completely at random (MCAR)
  - Missing at Random (MAR)
    - Measured variables, only, may influence missingness — including past Ys
  - Not Missing at Random (NMAR)
    - Depends on outcomes after dropout: really tough
- The distinctions matter because the type of missing data mechanism determines the analytic sophistication that is needed

Misspecified GEE
(when the truth is random intercepts and slopes)

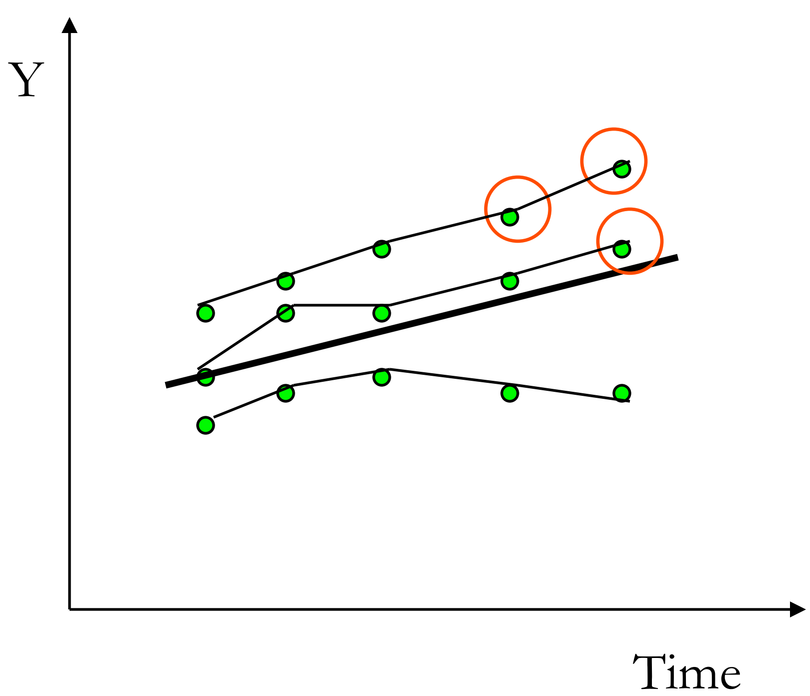
Complete Data (GEE)

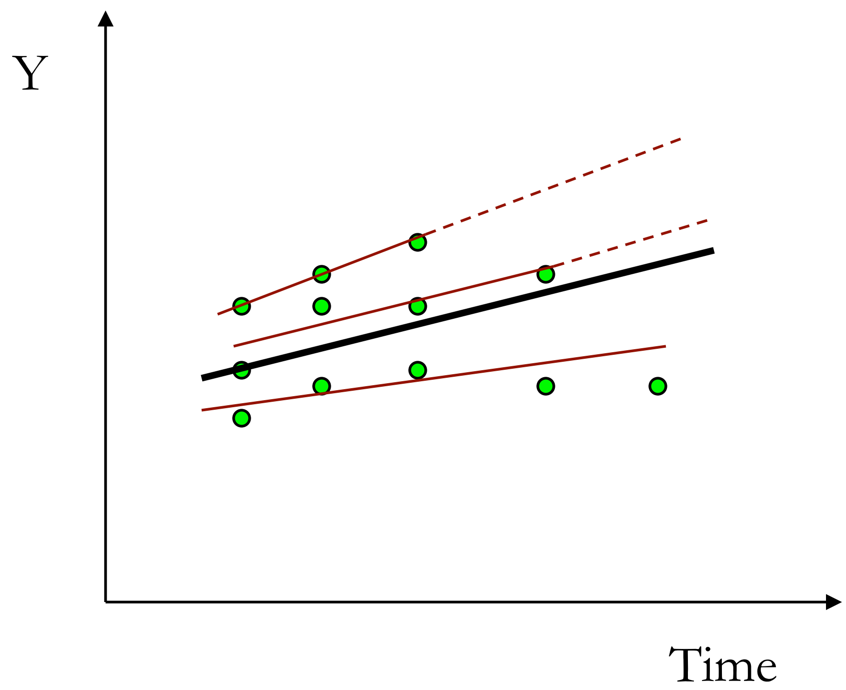


Partial Missing Data (GEE)



Correctly specified Random Effects
(when the truth is random intercepts and slopes)





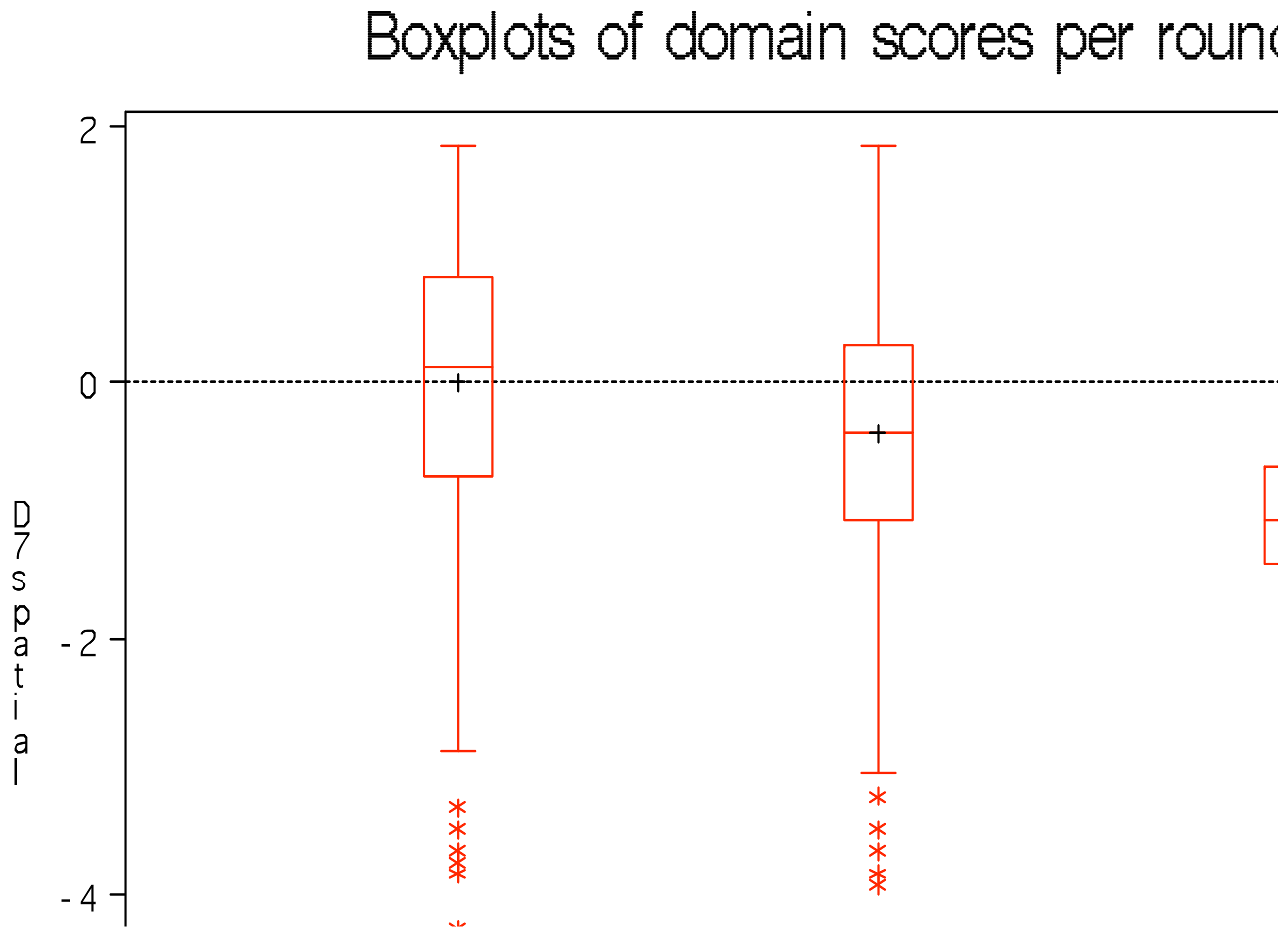
LDA Challenge # 3

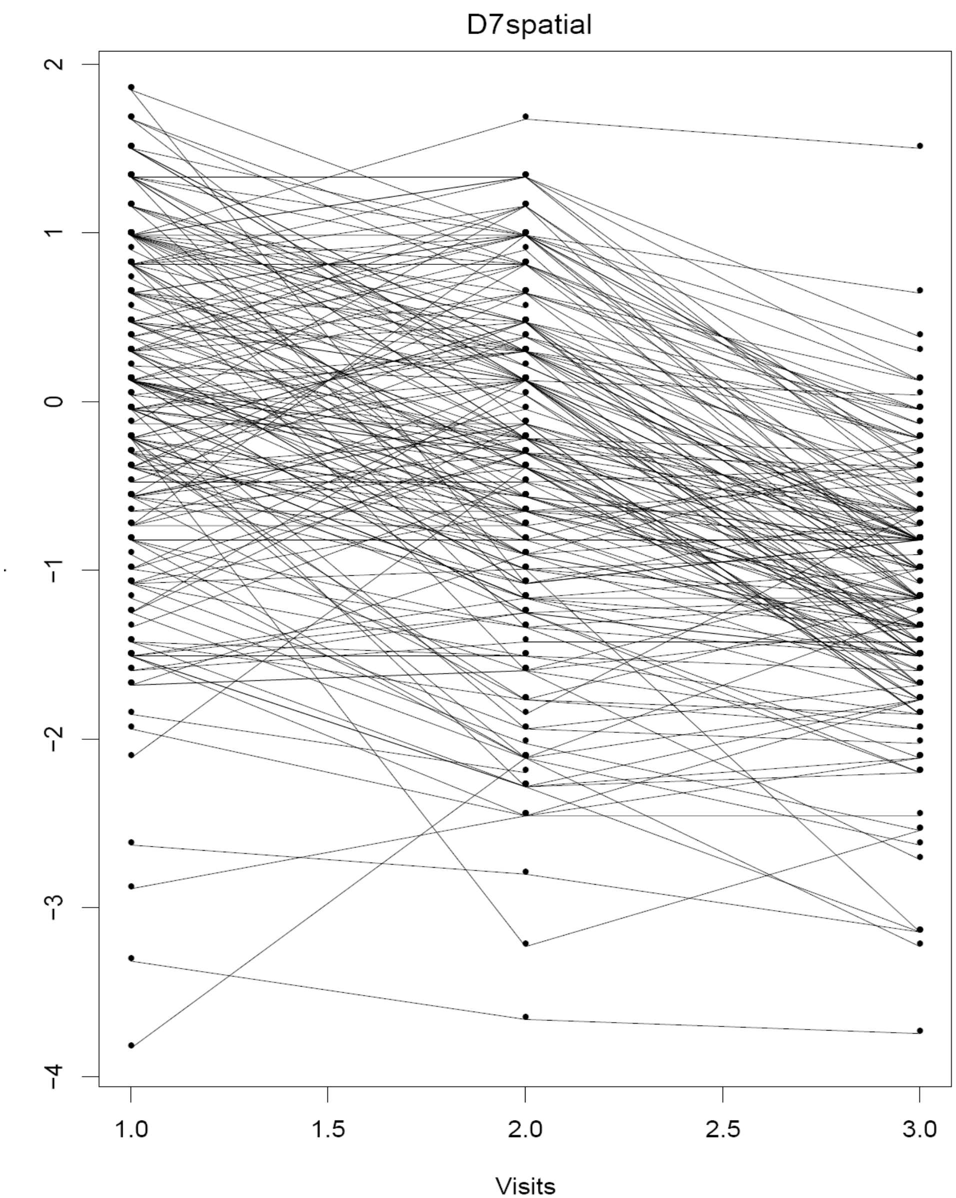
Nonlinear; clustered trajectories

A Second Example
Community Lead Exposure & Cognition

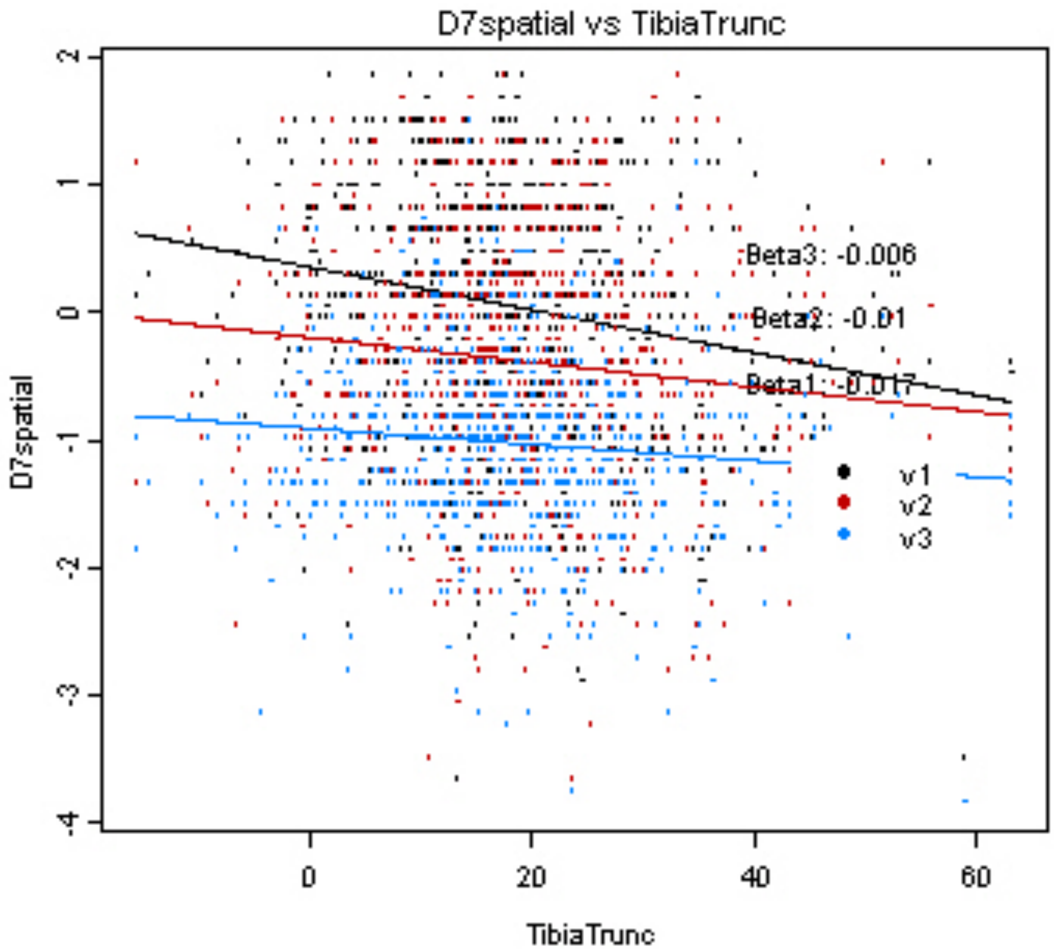
- Study: Baltimore Memory Study *(Schwartz et al., 2006)*
- Question: Does lead exposure affect visuo-spatial ability?
  - Tibia lead density (X: micrograms per gram)
    - A surrogate of lifetime dose
  - Visuo-spatial functioning (Y)
    - Z-score version of the Rey Copy test
  - Time (T)
    - Study rounds 0-2

Visuo-Spatial Scores in BMS
Side-by-side boxplots




D7spatial
2
1
0
−1
−2
−3
−4
1.0
1.5
2.0
2.5
3.0
Visits

Visuo-Spatial Scores in BMS Versus lead, by round



Take home points

- If you're out to save Millions at a Time©
  - Population average (marginal) model
    - Choice 1: GEE (corr-robust) vs. MLE (MAR-robust)
    - Choice 2: **Association structure to fit?**
  - Mean trajectory estimates not sensitive
- If one at a time, or seeking to target
  - Subject-specific (random effect) model
  - Benefit if model correct: heterogeneity characterization, MAR-robust, MLE: precise
- Temporality necessary, not sufficient, re causality
  - Transitions; time-varying covariates

- **It's all "Good."© Happy Modeling!**