Applying Group-Based Trajectory Modeling in Health Outcomes Research

Weihsuan ‘Jenny’ Lo-Ciganic, MSPharm, MS, PhD
Assistant Professor
Department of Pharmaceutical Outcomes & Policy, University of Florida
January 22, 2021 (12-2 PM EST)
Funding & Disclosures

- National Institute of Drug Abuse (R01DA044985)
- National Institute of Aging (R21 AG060308)
- National Institute of Mental Health (R01MH121907)
- PhRMA Foundation Research Starter Award
- No conflict of interests
Learning Objectives

• Overview the concept, capacities, and applications of group-based trajectory models (GBTM)
• Describe the key framework of conducting group-based trajectory models
• Understand the basic functions of group-based trajectory models using Stata
• Discuss the extensions and challenges of using GBTM
Disclaimer

➢ Learning everything about GBTM (even introduction-level) is too much for a 2-hour workshop

➢ Focus on concepts, applications and basic STATA tutorial

➢ Examples and emphasized focus on health and pharmaceutical outcomes related topics

➢ Omission is necessary in order to focus on the most important topics
At this point, you probably have many questions....

- How do you select the number of groups/trajectories? How do you evaluate model adequacy?
- How do you profile or describe group members?
- Can you add time-invariant covariates to the trajectory itself?
- Can you add time-varying covariates to the trajectory itself?
- Can you describe two or more behaviors/outcomes at the same time?
- What are differences between different methods to develop trajectory groups?
What is a trajectory?

A trajectory is “the evolution of an outcome over age or time.”

- GBTM was originally developed to study criminology and social behaviors (e.g., Montreal data in Nagin’s textbook)
An example where population-based average analysis fails

- Increasing (50%)
- Decreasing (50%)

Population Mean

Adherence measured by proportion of days covered (PDC)

Month
Motivations for using GBTM

➢ Test taxonomic theories

➢ Identify distinct development or behavioral paths from complex longitudinal data

➢ Provide more person-centered methods of analysis

➢ Summarize data with more transparency and visualized outputs

Important capabilities of GBTM

➢ Account for the dynamic outcome change over time
  o Medication utilization pattern changes can result from clinician’s decision, patient non-adherence, payer restrictions

➢ Identify differential patterns of individual change
  o Poorly identified by single annual adherence measure

➢ Characterize subgroups more likely to follow certain trajectories
  o Rather than arbitrarily assume or assign individuals to certain groups
  o Capable to estimate the proportion of the population following each trajectory

➢ Use groups to approximate an unknown distribution
  o Non-parametric or semi-parametric assumptions to allow flexibility

I. Overview of basic GBTM concepts

II. Applications in health and pharmaceutical outcomes research

III. Basic GBTM Methods with STATA tutorials

IV. Extensions and challenges of using GBTM
Association between Trajectories of Buprenorphine Treatment and Emergency Department and In-patient Utilization

Wei-Hsuan Lo-Ciganic\textsuperscript{1,2}, Walid F. Gellad\textsuperscript{2,3,4}, Adam J. Gordon\textsuperscript{2,3,4}, Gerald Cochran\textsuperscript{2,5}, Michael A. Zemaitis\textsuperscript{2,6}, Terri Cathers\textsuperscript{7}, David Kelley\textsuperscript{7} & Julie M. Donohue\textsuperscript{2,8}

### Rationale

- Little is known about current treatment patterns of buprenorphine for opioid use disorder.

### Question

- Is there a specific trajectory of buprenorphine use associated with adverse clinical outcomes?

### Methods

- A retrospective cohort study using 2007-2011 Pennsylvania Medicaid claims data

10,945 beneficiaries aged 18-64, non-dual eligible for Medicare who initiated buprenorphine fills

**Exposure:** (1) calculated interval-based *monthly proportion of days covered (PDC)* of buprenorphine for 1 year, and (2) used GBTM to identify buprenorphine trajectories

**Outcomes:** (1) time to first all-cause hospitalization, and (2) time to first emergency department visit

**Multivariable Cox proportional hazard models, adjusting for sociodemographics, health status and provider-level factors**

Overall PDC among Enrollees with Buprenorphine Prescriptions

Overall PDC over the index year: 0.44
Association between Trajectories of Buprenorphine Treatment and Emergency Department and inpatient Utilization

**All cause hospitalization**: compared to those who discontinued at 3-5 months
- Refilled persistently: 20% lower risk (HR=0.80, 95% CI: 0.68-0.94)

**ED visits**: compared to those who discontinued at 3-5 months
- Refilled persistently: 15% lower risk (HR=0.85, 95% CI: 0.77-0.94)
- Refilled intermittently: 21% higher risk (HR=1.21, 95% CI: 1.07-1.36)

Six Buprenorphine Trajectories among Pregnant Women with Opioid Use Disorder

73.9% initiated early
Who were the Early Adopters of Dabigatran? An Application of Group-based Trajectory Models


Figure 1. Trajectories of Physician Adoption of Dabigatran

- Rapid and extensive adopters (3.7%)
- Rapid and moderate adopters (13.4%)
- Slow adopters (21.6%)
- Minimal adopters (16.1%)
- Non-adopters (45.2%)
Association Between Dose and Duration Patterns of Opioid and Benzodiazepine Use and Risk of Overdose Among US Medicare Beneficiaries: A Group-based Multi-trajectory Model

Jenny Lo-Ciganic, PhD, Ting Wang, Yong Ge, Bobby L Jones, James Huang, Lili Zhou, Gary Reisfield, Jeannie K Lee, C. Kent Kowh, Juan M. Hincapie-Castillo, P. Chris Delcher, Khoa Nguyen, Chris Harle, Ching-Yuan Chang, Debbie L. Wilson, Jingchuan Guo, Walid F. Gellad

More details, see ICPE 2020 September 14 oral presentation (manuscript submitted)
Rationale and Scientific Question

➢ Rationale:
  o Concurrent opioid and benzodiazepine (OPI-BZD) use continues to rise despite clinical guidelines and US FDA black box warnings opposing such use.
  o Compared with younger adults, older adults have a greater prevalence of anxiety, insomnia and pain, and are 3 times more likely to be prescribed OPIs and BZDs.
  o The definitions of concurrent use vary substantially in the literature and have focused on arbitrary thresholds of duration (e.g., ≥1 day overlap) or dose alone.
  o Little is known about OPI-BZD dose and duration patterns most associated with OPI overdose risk.

➢ Question: What distinct dosing profiles of OPI-BZD use are associated with higher opioid overdose risk in Medicare?
Group-based Multi-Trajectory Models

➢ Advantage of group-based multi-trajectory models
  o Capture dynamic OPI and BZD dose changes simultaneously over time

Methods: study design and cohort

A retrospective cohort study using a 5% national sample of Medicare claims data (2013-2016)

Study cohort: 37,818 met the following inclusion/exclusion criteria

- Had ≥1 prescription fill for non-injectable, non-buprenorphine (for opioid use disorder) OPI or BZD
- Excluded:
  - Had a cancer diagnosis, were in hospice care or enrolled in Medicare Advantage plans
  - Did not have continuous enrollment 6 months prior to and after initiating OPIs
  - Had opioid or benzodiazepine overdose 6 months prior to and after initiating OPIs
  - Used only BZDs
  - Filled only 1 OPI or BZD prescription, or with <15 days of OPI or BZD supply during the 6-month trajectory measurement period (Pharmacy Quality Alliance criteria)
Methods: Statistical Analysis

Among 37,818 eligible Medicare beneficiaries

Step 1: Calculate average daily morphine milligram equivalent (MME) for OPIs and diazepam milligram equivalent (DME) for BZDs in the 6 months after initiating OPIs

Step 2: Identify distinct OPI-BZD dose and duration trajectories using group-based multi-trajectory model

Step 3: Calculate stabilized inverse probability of treatment weights (IPTW) for each beneficiary (*excluded extreme IPTW >10 [n=123])

Step 4: For each trajectory, estimate adjusted hazard ratios (HRs) of time to first OPI overdose episode within the 6 months following the 6-month trajectory measurement period using IPTW multivariable Cox model

(1) Austin PC. Multivariate Behav Res. 2011;46:399-424; (2) Faraone SV. P T. 2018;33:700-103, 710-711
Results: 9 OPI-BZD Trajectory Groups

Very-low-dose OPI (<25 MME)

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Labeling*</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>10,561</td>
<td>Very-low-dose OPI-BZD users with slowly decreasing BZD use (&lt;25 MME, &lt;10 DME)</td>
</tr>
<tr>
<td>B</td>
<td>4,900</td>
<td>Very-low-dose OPI-BZD users with consistent BZD use (&lt;25 MME, &lt;10 DME)</td>
</tr>
<tr>
<td>C</td>
<td>4,992</td>
<td>Very-low-dose OPI (&lt;25 MME) and medium-dose BZD (21-40 DME)</td>
</tr>
<tr>
<td>D</td>
<td>5,079</td>
<td>Low-dose OPI-BZD use (25-50 MME with 10-20 DME)</td>
</tr>
<tr>
<td>E</td>
<td>3,902</td>
<td>Low-dose OPI with high-dose BZD use (25-50 MME with 41-60 DME)</td>
</tr>
<tr>
<td>F</td>
<td>3,937</td>
<td>Medium-dose OPI with low-dose BZD use (51-100 MME with 10-20 DME)</td>
</tr>
<tr>
<td>G</td>
<td>1,360</td>
<td>Very-high-dose OPI with high-dose BZD use (&gt;150 MME with 41-60 DME)</td>
</tr>
<tr>
<td>H</td>
<td>949</td>
<td>Very-high-dose OPI with very-high-dose BZD use (&gt;150 MME with &gt;60 DME)</td>
</tr>
<tr>
<td>I</td>
<td>2,080</td>
<td>Very-high-dose OPI with low-dose BZD use (&gt;150 MME with 10-20 DME)</td>
</tr>
</tbody>
</table>

*Dose level in group labeling:
- Opioids: very low (<25 MME), low (25-50), moderate (51-90), high (91-150 MME), and very high (>150 MME).
- BZD dose level in diazepam equivalent milligram (DME): very low (<10 DME), low (10-20 DME), moderate (21-40 DME), high (41-60 DME), and very high (>60 DME).
Results: Dual Trajectories of OPI-BZD Use and Opioid Overdose Risk

<table>
<thead>
<tr>
<th>Group</th>
<th>% of the cohort</th>
<th>OPI overdose N (% of the trajectory)</th>
<th>Crude incidence (per 10,000 person mos)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>28.0</td>
<td>8 (0.08)</td>
<td>1.3</td>
</tr>
<tr>
<td>B</td>
<td>13.0</td>
<td>2 (0.04)</td>
<td>0.7</td>
</tr>
<tr>
<td>C</td>
<td>13.2</td>
<td>7 (0.14)</td>
<td>2.3</td>
</tr>
<tr>
<td>D</td>
<td>13.5</td>
<td>12 (0.24)</td>
<td>3.9</td>
</tr>
<tr>
<td>E</td>
<td>10.3</td>
<td>20 (0.51)</td>
<td>8.6</td>
</tr>
<tr>
<td>F</td>
<td>10.4</td>
<td>20 (0.51)</td>
<td>8.5</td>
</tr>
<tr>
<td>G</td>
<td>3.6</td>
<td>16 (1.18)</td>
<td>19.6</td>
</tr>
<tr>
<td>H</td>
<td>2.5</td>
<td>11 (1.16)</td>
<td>19.2</td>
</tr>
<tr>
<td>I</td>
<td>5.5</td>
<td>22 (1.06)</td>
<td>17.6</td>
</tr>
</tbody>
</table>

Groups E, and G to I accounted for 21.9% of the cohort and captured ~60% of opioid overdoses.
Conclusions

➢ 9 distinct OPI-BZD trajectories were identified during the 6 months following opioid initiation among Medicare fee-for-service beneficiaries

➢ OPI overdose risks varied substantially across OPI-BZD trajectories
  o Very-high-dose opioid use (MME >150) or high-dose benzodiazepine use (DME>40, even in the presence of low-dose opioid use) had a 2 to 4 times increased opioid overdose risk
  o 21% of the cohort were in the high-risk trajectories: captured ~60% of OPI overdoses

➢ Clinicians should avoid prescribing OPIs and BZDs concurrently whenever possible. When co-prescribing is necessary, clinicians should:
  o Discuss safety concerns with patients
  o Limit dosage and duration to the minimum required
  o Monitor closely with prescription drug monitoring program (PDMP)
Limitations

➢ Claim-based analyses have limited clinical and socio-behavior information such as pain severity
  - E-value ranged 3.6 to 6.7 for high-risk trajectory groups

➢ Unable to link to death certificate data and thus could not distinguish fatal from non-fatal overdoses

➢ Limited generalizability to other populations (e.g., Medicaid)

➢ Unable to evaluate the impact of US FDA black box warning released in August 2016

Dual-Trajectories of Opioid and Gabapentinoid Use and Risk of Subsequent Drug Overdose among Medicare Beneficiaries in the US

Dual-trajectories of opioid and gabapentinoid use and risk of subsequent drug overdose among Medicare beneficiaries in the United States: a retrospective cohort study

Lili Zhou¹, Sandipan Bhattacharjee¹, C. Kent Kwoh²,³, Patrick J. Tighe⁴, Gary M. Reisfield⁵, Daniel C. Malone⁶, Marion Slack¹, Debbie L. Wilson⁷, Ching-Yuan Chang⁷,⁸ & Wei-Hsuan Lo-Ciganic⁷,⁸

More details, see Addiction 2020 Jul 10.
Key Results & Main Conclusions

- Subsequent overdose risk varied substantially by different OPI-GABA trajectories.
- High-dose OPI-only users and all consistent OPI-GABA users were associated with more than doubled drug overdose risk.

A. OPI only (No GABA): 59.0%
   - OPI-only early discontinuers (40.6%)
   - Reference

B. Consistent low-dose OPI-only users (16.6%)
   - HR= 1.47 (1.19, 1.82)
   - 4.57 (2.99, 6.98)

C. Consistent high-dose OPI-only users (1.8%)
   - Early discontinuation of OPIs and consistent low-dose GABA (6.9%)

D. GABA only (No OPI use): 26.6%
   - GABA-only early discontinuers (12.5%)
   - 1.39 (1.09, 1.77)

E. Consistent low-dose GABA-only users (11.0%)
   - 1.44 (1.12, 1.85)

F. Consistent high-dose GABA-only users (3.1%)
   - 1.43 (0.94, 2.17)

G. Concurrent OPI-GABA: 14.4%
   - Early discontinuation of OPIs and consistent low-dose GABA (6.9%)

H. Consistent low-dose OPI-GABA users (3.4%)
   - Consistent low-dose OPI and high-dose GABA users (3.2%)

I. Consistent low-dose OPI and moderate-dose GABA users (0.9%)

J. Consistent high-dose OPI and moderate-dose GABA users (0.9%)

- Reference HR= 1.47 (1.19, 1.82)
- 4.57 (2.99, 6.98)
- 1.24 (0.90, 1.69)
- 2.49 (1.76, 3.52)
- 2.46 (1.71, 3.53)
- 7.22 (4.46, 11.7)
Dual-Trajectories of Opioid and Gabapentinoid Use and Health Expenditures among Medicare Beneficiaries in the US

Association Between Dual Trajectories of Opioid and Gabapentinoid Use and Healthcare Expenditures Among US Medicare Beneficiaries

Lili Zhou, PhD, Sandipan Bhattacharjee, PhD, C. Kent Kwoh, MD, Daniel C. Malone, PhD, Patrick J. Tighe, MD, Gary M. Reisfield, MD, Marion Slack, PhD, Debbie L. Wilson, PhD, Wei-Hsuan Lo-Ciganic, PhD

More details, see Value in Health 2021 (in press)
<table>
<thead>
<tr>
<th>Group</th>
<th>Mean total annual concurrent direct medical costs (95% CI)</th>
<th>Cost ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>$13,830 ($13,643-$14,019)</td>
<td>Reference</td>
</tr>
<tr>
<td>B</td>
<td>$15,721 ($15,395-$16,055)</td>
<td>1.14 (1.11-1.17)</td>
</tr>
<tr>
<td>C</td>
<td>$22,908 ($21,421-$24,497)</td>
<td>1.66 (1.55-1.77)</td>
</tr>
<tr>
<td>D</td>
<td>$10,607 ($10,345-$10,876)</td>
<td>0.77 (0.75-0.79)</td>
</tr>
<tr>
<td>E</td>
<td>$12,397 ($12,053-$12,751)</td>
<td>0.89 (0.87-0.92)</td>
</tr>
<tr>
<td>F</td>
<td>$11,713 ($11,254-$12,191)</td>
<td>0.85 (0.81-0.88)</td>
</tr>
<tr>
<td>G</td>
<td>$13,659 ($12,574-$14,838)</td>
<td>0.99 (0.91-1.07)</td>
</tr>
<tr>
<td>H</td>
<td>$18,309 ($17,743-$18,893)</td>
<td>1.32 (1.28-1.37)</td>
</tr>
<tr>
<td>I</td>
<td>$22,869 ($21,841-$23,946)</td>
<td>1.65 (1.58-1.73)</td>
</tr>
<tr>
<td>J</td>
<td>$20,281 ($19,211-$21,411)</td>
<td>1.47 (1.39-1.55)</td>
</tr>
<tr>
<td>K</td>
<td>$28,464 ($25,910-$31,271)</td>
<td>2.06 (1.87-2.26)</td>
</tr>
</tbody>
</table>
I. Overview of basic GBTM concepts

II. Applications in health and pharmaceutical outcomes research

III. Basic GBTM Methods with STATA tutorials

IV. Extensions and challenges of using GBTM
GBTM specifications: maximum likelihood estimation

\[ P(Y_i) = \sum_j \pi_j(X_i) P^j(Y_i) \]

\[ \pi_j(X_i) = \frac{e^{X_i\theta_j}}{\sum e^{X_i\theta_j}} \]

\[ \prod_{i=1}^{N} P(Y_i) \]

\( Y_i \): longitudinal patterns of outcomes
\( j \): group or trajectory \( j \)
\( \pi_j \): probability of membership in group \( j \)
\( X_i \): covariates
\( P^j(Y_i) \): probability of \( Y_i \) given membership in group \( j \)

### Types of Outcome Data in GBTM

\[ P(Y_i) = \sum_j \pi_j (X_i) P^j(Y_i) \]

- **Continuous data**: Medication adherence measured by proportion of days covered (PDC), psychometric scale
  - Distribution/model: Censored normal model or beta-distributed model
- **Count data**: Number of readmission, number of times in jail
  - Distribution/model: Poisson-based model (e.g., zero-inflated Poisson [ZIP])
- **Binary data**: PDC >80% (yes vs. no)
  - Distribution/model: Logit-based model

\[ Y_i: \text{longitudinal patterns of outcomes} \]
Software for GBTM

- Free and easy to use: SAS and STATA-based procedure
  - Available for download at [https://www.andrew.cmu.edu/user/bjones/](https://www.andrew.cmu.edu/user/bjones/)

- Provide confidence intervals on trajectory estimates

- Accommodate missing data

- Can handle sample weights (e.g., for panel data)

- Allow for irregular time spacing of measurement

- Accommodate over-lapping cohort designs
Traj in STATA

➢ https://www.andrew.cmu.edu/user/bjones/

➢ To install the Stata version:
  net from http://www.andrew.cmu.edu/user/bjones/traj
  net install traj, force
  help traj

Examples
1. Censored normal (cnorm) model
2. Variability (sigma) by group option - cnorm model
3. Zero-inflated Poisson (zip) model
4. Logistic (logit) model
5. Providing start values
6. Including time-stable covariates (risk) associated with group membership
7. Obtaining group membership probabilities from a model with risk variables
8. Including covariates (tcov), specified at each time point, associated with group trajectory means
9. Parametric bootstrap sampling for model parameters e.g. group size confidence intervals
10. Wald hypothesis tests of the traj model parameters (like SAS %trajtest)
11. Distal outcome model
12. Distal outcome model with bootstrap CI for individual outcome predictions
13. Joint trajectory model
14. Multi-trajectory model
15. Dropout modeling
16. Exposure time / sample weights
Traj: basic syntax

```
traj [if], var(varlist) indep(varlist) model(modeltype) order(numlist) [additional options]
```

---

<table>
<thead>
<tr>
<th>options</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Trajectory Variables</strong></td>
<td></td>
</tr>
<tr>
<td>var(varlist)</td>
<td>dependent variables, measured at different times or ages</td>
</tr>
<tr>
<td>indep(varlist)</td>
<td>independent variables i.e. when the dependent variables were measured</td>
</tr>
<tr>
<td><strong>Model</strong></td>
<td></td>
</tr>
<tr>
<td>model(modeltype)</td>
<td>beta, cnorm (censored normal), logit (Hosmer-Lemeshow), zip (zero-inflated Poisson) - probability distribution for the dependent variables</td>
</tr>
<tr>
<td>order(numlist)</td>
<td>@intercept, 1linear, 2quadratic, 3cubic - polynomial type for each group trajectory</td>
</tr>
<tr>
<td>min()</td>
<td>(cnorm, default to 0) minimum value for the censored normal model</td>
</tr>
<tr>
<td>max()</td>
<td>(required for cnorm) maximum value for the censored normal model</td>
</tr>
<tr>
<td>sigma(bygroup)</td>
<td>optional polynomial type (0-intercept, 1-linear, 2-quadratic, 3-cubic) for the zero-inflation of each group</td>
</tr>
<tr>
<td>exposure(varlist)</td>
<td>optional exposure variables for the zero-inflated Poisson model</td>
</tr>
<tr>
<td>weight(varname)</td>
<td>optional sampling weight variable</td>
</tr>
<tr>
<td><strong>Time-Stable Covariates For Group Membership</strong></td>
<td></td>
</tr>
<tr>
<td>ref(group)</td>
<td>controls the reference group (default = 1) when the risk option is used</td>
</tr>
<tr>
<td><strong>Time-Varying Covariates Influencing Trajectory Paths</strong></td>
<td></td>
</tr>
<tr>
<td>tcv(varlist)</td>
<td>time-varying covariates for each group trajectory</td>
</tr>
<tr>
<td>plotcv(varlist)</td>
<td>optional values for plotting trajectories with time-varying covariates</td>
</tr>
<tr>
<td><strong>Dropout Model</strong></td>
<td></td>
</tr>
<tr>
<td>dropout(varlist)</td>
<td>include logistic model of dropout probability per wave with (\theta) - constant rate, (1) - depends on the previous response, (2) - depends on the two previous responses, for each group</td>
</tr>
<tr>
<td>dcv(varlist)</td>
<td>optional lagged time-varying covariates for the dropout model</td>
</tr>
<tr>
<td>observer(varname)</td>
<td>optional binary variable to mark which observations are to be included in the dropout model and those to be treated as missing at random. This variable = 1 for observations to be treated as data MAR (include completers) and = 0 for observations to be used for the modeled dropout</td>
</tr>
<tr>
<td><strong>Distal Outcome Model</strong></td>
<td></td>
</tr>
<tr>
<td>outcome(varname)</td>
<td>a distal variable to be regressed on the probability of group membership</td>
</tr>
<tr>
<td>omode(modeltype)</td>
<td>probability distribution for the outcome variable: normal, logit, mlogit, or poisson</td>
</tr>
<tr>
<td>omc(varlist)</td>
<td>optional covariates for the outcome model</td>
</tr>
<tr>
<td><strong>Joint Trajectory Model</strong></td>
<td></td>
</tr>
<tr>
<td>The joint trajectory model uses the options shown above with a 'J' suffix to specify the second model e.g. model1(modeltype) etc. See the joint trajectory example.</td>
<td></td>
</tr>
<tr>
<td><strong>Multi-Trajectory Model</strong></td>
<td></td>
</tr>
<tr>
<td>multi(varlist)</td>
<td>the number of multi-trajectory group memberships (2 to 6). The multi-trajectory model uses the options shown above with a 'J', 'S', etc. (up to 6) suffix to specify the additional models. See the multi-trajectory example.</td>
</tr>
<tr>
<td><strong>Other</strong></td>
<td></td>
</tr>
<tr>
<td>start(varlist)</td>
<td>parameter start values to override default start values</td>
</tr>
<tr>
<td>detail</td>
<td>shows start values, maximization iterations, and ending values for monitoring model fitting progress. The ending values can be useful as start values for future traj models.</td>
</tr>
<tr>
<td>ci</td>
<td>parametric bootstrap confidence intervals of individual distal outcome and probability of group memberships.</td>
</tr>
<tr>
<td>reps</td>
<td>number of bootstrap replications (default = 1000).</td>
</tr>
<tr>
<td>scoreci</td>
<td>confidence intervals of individual distal outcome and probability of group memberships using the method of Sison and Glaz (1995).</td>
</tr>
</tbody>
</table>
Basic Data Layout for GBTM

### Var(varlist): longitudinal outcome patterns of interest (dependent variable in GBTM)

The table below shows the basic data layout for GBTM, with Patient_IDs from 001 to 007. Each patient has multiple measurements over time, with specific varlist values denoted as PDC_1, PDC_2, PDC_3, ..., PDC_12. The time points for these measurements are indicated in columns labeled Time_1, Time_2, ..., Time_12.

### Indep(varlist): age or time when dependent variables were measured

- **Patient_ID**: Unique identifier for each patient.
- **PDC_1** to **PDC_12**: Variables representing different outcome patterns of interest for each patient.
- **Time_1** to **Time_12**: Time points at which these variables were measured for each patient.

| Patient_ID | PDC_1 | PDC_2 | PDC_3 | PDC_4 | PDC_5 | PDC_6 | PDC_7 | PDC_8 | PDC_9 | PDC_10 | PDC_11 | PDC_12 | Time_1 | Time_2 | Time_3 | Time_4 | Time_5 | Time_6 | Time_7 | Time_8 | Time_9 | Time_10 | Time_11 | Time_12 |
|------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 001        | 1.00  | 0.90  | 0.75  | 0.55  |       |       |       |       |       |       |       |       | 1     | 2     | 12    |       |       |       |       |       |       |       |       |       |       |
| 002        | 0.85  | 0.75  | 0.55  | 0.40  |       |       |       |       |       |       |       |       | 1     | 2     | 12    |       |       |       |       |       |       |       |       |       |       |
| 003        | 0.50  | 0.35  | 0.25  | 0.55  |       |       |       |       |       |       |       |       | 1     | 2     | 12    |       |       |       |       |       |       |       |       |       |       |
| 004        | 0.35  | 0.55  | 0.65  | 0.75  |       |       |       |       |       |       |       |       | 1     | 2     | 12    |       |       |       |       |       |       |       |       |       |       |
| 005        | 0.75  | 0.55  | 0.65  | 0.75  |       |       |       |       |       |       |       |       | 1     | 2     | 12    |       |       |       |       |       |       |       |       |       |       |
| 006        | 0.77  | 0.80  | 0.82  | 0.85  |       |       |       |       |       |       |       |       | 1     | 2     | 12    |       |       |       |       |       |       |       |       |       |       |
| 007        | 1.00  | 0.00  | 0.50  | 0.75  |       |       |       |       |       |       |       |       | 1     | 2     | 12    |       |       |       |       |       |       |       |       |       |       |
| ...        | ...   | ...   | ...   | ...   |       |       |       |       |       |       |       |       | ...   | ...   | ...   |       |       |       |       |       |       |       |       |       |       |

Note: The values for Patient_IDs 008 to 007 are not shown as they follow the same pattern.
**GBTM: Censored normal (Tobit) model**

**Censored normal model:**

\[
y_{it}^j = \beta_0^j + \beta_1^j \times age_{it} + \beta_2^j \times age_{it}^2 + \beta_3^j \times age_{it}^3 + \varepsilon
\]

- \(j\): group/number of groups

**Time polynomial order in GBTM:**
- **liner** = \(time^1\)
- **quadratic** = \(time^2\)
- **cubic** = \(time^3\)
- **quartic** = \(time^4\)
- **quintic** = \(time^5\)

**For example:**
\[
y^* = \beta_0 + \beta_1 \times age + \varepsilon
\]

**Definition of censored normal distribution**

- \(y = 0\) if \(y^* \leq 0\)
- \(y = y^*\) if \(0 < y \leq y^{\max}\)
- \(y^* = y^{\max}\) if \(y^* > y^{\max}\)
STATA: Censored normal model example

Montreal data: The data consist of annual assessments on 1,037 boys at age 6 (spring 1984) and ages 10 through 15 on an oppositional behavior scale (ranges from 0 to 10) gathered in low socioeconomic areas of Montreal, Canada. See Tremblay et al. (1987) for details. Scores of zero are frequent and the scores decrease in frequency as the score increases. Hence, the censored normal distribution is sensible for modeling the data. The following commands fit a 3-group model to the opposition data and provide a graph of the results.

STATA Syntax

use http://www.andrew.cmu.edu/user/bjones/traj/data/montreal_sim.dta, clear
traj, var(qcp*op) indep(age*) model(cnorm) min(0) max(10) order(1 3 2)
trajplot, xtitle(Age) ytitle(Opposition) xlabel(6(1)15) ylabel(0(1)6)
list _traj_Group - _traj_ProbG3 if _n < 3, ab(12)
matrix list e(plot1), format(%9.2f) noheader

### STATA: Censored normal model output (Example 1)

**1037 observations used in the trajectory model.**

| Group | Parameter  | Estimate | Error | T for H0: Parameter=0 | Prob > |T| |
|-------|------------|----------|-------|------------------------|--------|---|
| 1     | Intercept  | 1.00232  | 0.35105 | 2.855                  | 0.0043 | |
|       | Linear     | -0.19098 | 0.03065 | -6.232                 | 0.0000 | |
| 2     | Intercept  | -13.84777| 4.09008 | -3.386                 | 0.0007 | |
|       | Linear     | 4.95268  | 1.30882 | 3.786                  | 0.0002 | |
|       | Quadratic  | -0.45532 | 0.12967 | -3.511                 | 0.0004 | |
|       | Cubic      | 0.01261  | 0.00407 | 3.094                  | 0.0020 | |
| 3     | Intercept  | -1.61687 | 0.97421 | -1.659                 | 0.0972 | |
|       | Linear     | 1.45292  | 0.19585 | 7.418                  | 0.0000 | |
|       | Quadratic  | -0.07251 | 0.00393 | -7.721                 | 0.0000 | |
|       | Sigma      | 2.61114  | 0.03276 | 79.709                 | 0.0000 | |

**Group membership**

<table>
<thead>
<tr>
<th>Group</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>30.84358</td>
</tr>
<tr>
<td>2</td>
<td>46.12672</td>
</tr>
<tr>
<td>3</td>
<td>23.02970</td>
</tr>
</tbody>
</table>

**BIC=-11908.18 (N=6231)  BIC=-11897.42 (N=1037)  AIC=-11867.75  ll= -11855.75**

---

**Model estimates (95%CI)**

- **Group 1:** 30.8% (95% CI: 29.3% - 32.3%)
- **Group 2:** 46.1% (95% CI: 44.6% - 47.6%)
- **Group 3:** 23.0% (95% CI: 21.5% - 24.5%)
GBTM: Poisson-based model

Basic Poisson-based model: $\log(\lambda_{it}) = \beta_0^i + \beta_1^i \times age_{it} + \beta_2^i \times age_{it}^2 + \beta_3^i \times age_{it}^3$

$\lambda$: mean value (e.g., event rate)

Zero – inflated Posissson Model: using when there are a lot of 0s in Poisson model

$$p(x) = \begin{cases} 0 \text{ with probability } \rho \\ \text{Poisson}(\lambda) \text{ with probability } 1 - \rho \end{cases}$$

$$\ln(\lambda) = \beta_0 + \beta_1 \times age + \beta_2 \times age^2 + \beta_3 \times age^3$$

$$\rho = \frac{e^{\alpha_0 + \alpha_1 \times age + \alpha_2 \times age^2 + \alpha_3 \times age^3}}{1 + e^{\alpha_0 + \alpha_1 \times age + \alpha_2 \times age^2 + \alpha_3 \times age^3}}$$
The data are the annual number of criminal offense convictions for 411 subjects from a prospective longitudinal survey conducted in a working-class section of London (Farrington and West, 1990). The annual criminal offense convictions were recorded for boys from age 10 through age 30. The Poisson model is appropriate here; however, more zeros are present than would be expected in the purely Poisson model, so we will use the ZIP model. The following commands fit a 3-group model to the data and provide a graph of the results.

**STATA Syntax**

```stata
use http://www.andrew.cmu.edu/user/bjones/traj/data/anag1.dta, clear
traj, var(y*) indep(t*) model(zip) order(0 3 3) iorder(0 -1 0)
/* t1-t11 were scaled from -1 to 1 that may work with the default start values, but no guarantee sometimes */
/* The following Stata commands return the x-axis to the original time scale. */
mat P = e(plot1)
svmat P, names(col)
replace trajT = 10 * trajT + 40
trajplot, xtitle(Age) ytitle(Annual Conviction Rate) plotvars(trajT-U953) ci
drop trajT - U953
/* Assigned group and probabilities of group membership */
list _traj_Group - _traj_ProbG3 if _n > 400, ab(12)
```

STATA: ZIP model output (Example 3)

403 observations read.
6 had no trajectory data.
397 observations used in the trajectory model.

Maximum Likelihood Estimates
Model: Zero Inflated Poisson (zip)

| Group | Parameter | Estimate | Standard Error | T for H0: Parameter=0 | Prob > |T| |
|-------|-----------|----------|----------------|------------------------|--------|---|
| 1     | Intercept | -1.41663 | 0.84311        | -1.680                 | 0.0930 |
| 2     | Intercept | -1.29319 | 0.23242        | -5.564                 | 0.0000 |
|       | Linear    | -2.46477 | 0.55400        | -4.449                 | 0.0000 |
|       | Quadratic | -1.93476 | 0.40930        | -4.727                 | 0.0000 |
|       | Cubic     | 2.88839  | 0.78768        | 3.667                  | 0.0002 |
| 3     | Intercept | 0.61512  | 0.09540        | 6.448                  | 0.0000 |
|       | Linear    | -1.09516 | 0.24499        | -4.470                 | 0.0000 |
|       | Quadratic | -1.25101 | 0.19201        | -6.515                 | 0.0000 |
|       | Cubic     | 1.44852  | 0.35736        | 4.053                  | 0.0001 |
| 1     | Alpha0    | 3.05252  | 0.76850        | 3.972                  | 0.0001 |
| 3     | Alpha0    | -0.72287 | 0.22094        | -3.272                 | 0.0011 |

Group membership

<table>
<thead>
<tr>
<th>Group</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>68.26470</td>
</tr>
<tr>
<td>2</td>
<td>20.19498</td>
</tr>
<tr>
<td>3</td>
<td>11.54033</td>
</tr>
</tbody>
</table>

BIC = -1491.66 (N=4367)  BIC = -1476.07 (N=397)  AIC = -1450.17  ll = -1437.17

![Annual Conviction Rate vs Age graph]
GBTM: Logistic (logit) model

\[ p(y = 1) = \frac{e^{\beta_0 + \beta_1 \times \text{age} + \beta_2 \times \text{age}^2 + \beta_3 \times \text{age}^3}}{1 + e^{\beta_0 + \beta_1 \times \text{age} + \beta_2 \times \text{age}^2 + \beta_3 \times \text{age}^3}} \]

where \( \begin{cases} y = 1 \text{ if yes} \\ y = 0 \text{ if no} \end{cases} \)
It is common in research on criminal careers to analyze the absence or presence of offenses (i.e. a dichotomous prevalence measure). The ZIP analysis is repeated for a derived criminal offense prevalence measure using a logistic model (i.e., periods in which 1 or more convictions are reported are coded as “1” and periods with no convictions are coded as “0”). The following commands fit a three-group model to the prevalence measure data and graph the results.

```
use http://www.andrew.cmu.edu/user/bjones/traj/data/cambrdge.dta, clear
traj, var(p1-p23) indep(tt1-tt23) model(logit) order(0 3 3)
trajplot, xtitle(Scaled Age) ytitle(probability of presence of offenses)
/* Assigned group and probabilities of group membership */
list _traj_Group - _traj_ProbG3 if _n > 400, ab(12)
```
The logistic model gives the log-odds of response. The log-odds is converted to the probability of response.
GBTM: Beta-distribution based trajectory

- An alternative to the normal distribution for modeling continuous longitudinal data that are poorly fit by the normal distribution even with censoring.
- Primary advantage: the flexibility of the shape of the density function
- Disadvantage: the data under study must be transformable to a 0–1 scale.

Fig. 1 The Distribution of Hour 12 Suppression Ratio Data with the Best Fitting Beta Distribution. *The sum of the heights of the relative frequency density bars multiplied by their width sum to 1.0 so as to conform the with estimated beta density.

Posterior Probability of Group Membership (PPGM)

\[ PPGM = AvePP_j = \hat{p}(\text{group } j | \text{data}_i) = \frac{\hat{p}(\text{data}_i | \text{group } j)\hat{\pi}_j}{\sum_j \hat{p}(\text{data}_i | \text{group } j)\hat{\pi}_j} \]

\[ \hat{p}(\text{data}_i | \text{group } j) \]: probability of your data, given group membership
\[ \hat{\pi}_j \]: probability of being in group j

➢ Maximum probability group assignment rule: Bayes’ rule
  o Used to assign individual to group j in which they have the largest posterior probability

➢ Other uses of PPGM: one of the most important values/features in GBTM
  o Diagnostics for model fit (i.e., PPGM >0.7)
  o Match people with comparable developmental histories (e.g., used with propensity score)
  o Compute weighted averages that account for group membership uncertainty
  o Can be further used as serial measured to examine how quickly you can correctly estimate which trajectory an individual will ultimately follow

Outline

I. Overview of basic GBTM concepts

II. Applications in health and pharmaceutical outcomes research

III. Basic GBTM Methods with STATA tutorials

IV. Extensions and challenges of using GBTM
At this point, you probably have many questions….

How do you select the number of groups/trajectories? How do you evaluate model adequacy?

How do you profile or describe group members?

Can you add time-invariant covariates to the trajectory itself?

Can you add time-varying covariates to the trajectory itself?

Can you describe two or more behaviors/outcomes at the same time?

What are differences between different methods to develop trajectory groups?
Select GBTM final model & model evaluation

➢ “Forward” classifying approach (i.e., adding one extra group at a time)
➢ A combination of BIC and Nagin’s criteria
   o BIC: larger is better (i.e., BIC more towards to right direction of x-axis is better!)
     \[ BIC = \log(L) - 0.5 \times \log(n) \times k \]
     (where \( L \): log likelihood, \( n \): sample size, \( k \): number of parameters)
   o Nagin’s criteria
     ❑ Average posterior probability of assignment (PPGM or AvePP\(_j\)) for all J groups >0.7
     ❑ Odds of correct classification (OCC\(_j\)) \( \geq 0.5 \), where \( OCC_j = \frac{\text{AvePP}_j}{1 - \text{AvePP}_j} \)
     ❑ Model estimate (\( \hat{\pi}_j \)) close to proportion of sample assigned to j (\( \frac{N_j}{N} \))
     ❑ Confidence intervals for \( \hat{\pi}_j \) reasonably narrow
➢ Challenge: BIC keeps improving when number of groups increases, especially using large data→ usually we cannot only rely on BIC

### Appendix Table 1. Bayesian Information Criterion (BIC) Values and Predicted Group Proportions for Group-Based Trajectory Models for 2-, 3-, 4-, 5-, 6-, and 7-Group Trajectory Solutions

<table>
<thead>
<tr>
<th>Groups</th>
<th>No. of Groups</th>
<th>BIC&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Group 1</th>
<th>Group 2</th>
<th>Group 3</th>
<th>Group 4</th>
<th>Group 5</th>
<th>Group 6</th>
<th>Group 7</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td></td>
<td>-102214.5</td>
<td>54.8%</td>
<td>45.2%</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>-94942.81</td>
<td>40.2%</td>
<td>24.9%</td>
<td>34.9%</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td>-91482.74</td>
<td>37.8%</td>
<td>9.4%</td>
<td>23.4%</td>
<td>29.3%</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>-87573.36</td>
<td>29.3%</td>
<td>9.3%</td>
<td>20.8%</td>
<td>16.6%</td>
<td>23.9%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td>-86246.70</td>
<td>9.5%</td>
<td>24.9%</td>
<td>12.3%</td>
<td>13.3%</td>
<td>18.7%</td>
<td>21.2%</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>-89498.81</td>
<td>14.2%</td>
<td>13.8%</td>
<td>14.3%</td>
<td>14.3%</td>
<td>14.3%</td>
<td>14.8%</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviation:** BIC: Bayesian information criterion

<sup>a</sup> This value is based on the model likelihood with a penalty for the number of model parameters. It is not directly interpretable from the table, but the higher value here indicates better model fit.

### Appendix Table 2. Final 6-Group Group-Based Trajectory Model for Buprenorphine Refill Patterns

<table>
<thead>
<tr>
<th>Group (Pattern)</th>
<th>Estimated (95% Confidence Interval)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discontinued at 3 months</td>
<td>Intercept: 1.75 (1.70, 1.80)</td>
<td>72.06</td>
</tr>
<tr>
<td>Discontinued at 3 months</td>
<td>Month: -0.01 (0.04, -0.08)</td>
<td>-46.58</td>
</tr>
<tr>
<td>Discontinued at 3 months</td>
<td>Month&lt;sup&gt;b&lt;/sup&gt;: 0.11 (0.10, 0.12)</td>
<td>25.70</td>
</tr>
<tr>
<td>Discontinued at 3 months</td>
<td>Month&lt;sup&gt;c&lt;/sup&gt;: -0.0037 (-0.0041, -0.0032)</td>
<td>16.01</td>
</tr>
<tr>
<td>Discontinued at 5-8 months</td>
<td>Intercept: 1.42 (1.34, 1.51)</td>
<td>32.85</td>
</tr>
<tr>
<td>Discontinued at 5-8 months</td>
<td>Month: -0.41 (0.45, -0.33)</td>
<td>-10.29</td>
</tr>
<tr>
<td>Discontinued at 5-8 months</td>
<td>Month&lt;sup&gt;b&lt;/sup&gt;: 0.13 (0.11, 0.15)</td>
<td>12.37</td>
</tr>
<tr>
<td>Discontinued at 5-8 months</td>
<td>Month&lt;sup&gt;c&lt;/sup&gt;: -0.013 (0.014, -0.011)</td>
<td>-18.77</td>
</tr>
<tr>
<td>Discontinued after 8 months</td>
<td>Intercept: 1.31 (1.25, 1.37)</td>
<td>42.93</td>
</tr>
<tr>
<td>Discontinued after 8 months</td>
<td>Month: -0.18 (0.22, -0.14)</td>
<td>-8.55</td>
</tr>
<tr>
<td>Discontinued after 8 months</td>
<td>Month&lt;sup&gt;b&lt;/sup&gt;: 0.045 (0.037, 0.053)</td>
<td>11.13</td>
</tr>
<tr>
<td>Discontinued after 8 months</td>
<td>Month&lt;sup&gt;c&lt;/sup&gt;: -0.0037 (-0.0042, -0.0033)</td>
<td>-17.17</td>
</tr>
</tbody>
</table>

### Appendix Table 3. Nagin’s Diagnostic Criteria for Group-Based Trajectory Model

<table>
<thead>
<tr>
<th>Group (Pattern)</th>
<th>Model Estimate of Group Probability (95% CI)&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Proportion Classified in Group&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Average Posterior Probability&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Odds Correct Classification&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discontinued at 1-3 months</td>
<td>0.249 (0.239, 0.259)</td>
<td>0.249</td>
<td>0.94</td>
<td>47.33</td>
</tr>
<tr>
<td>Discontinued at 3-5 months</td>
<td>0.187 (0.177, 0.197)</td>
<td>0.188</td>
<td>0.90</td>
<td>38.98</td>
</tr>
<tr>
<td>Discontinued at 5-8 months</td>
<td>0.123 (0.116, 0.131)</td>
<td>0.126</td>
<td>0.91</td>
<td>70.43</td>
</tr>
<tr>
<td>Discontinued after 8 months</td>
<td>0.133 (0.125, 0.141)</td>
<td>0.130</td>
<td>0.92</td>
<td>77.14</td>
</tr>
<tr>
<td>Refilled intermittently</td>
<td>0.095 (0.089, 0.101)</td>
<td>0.095</td>
<td>0.93</td>
<td>126.67</td>
</tr>
<tr>
<td>Refilled persistently</td>
<td>0.212 (0.204, 0.221)</td>
<td>0.214</td>
<td>0.96</td>
<td>88.40</td>
</tr>
</tbody>
</table>

<sup>a</sup> 95% confidence intervals (CIs), based on parametric bootstrap method, should be reasonably narrow.

<sup>b</sup> Proportion classified in group is based on the maximum posterior probability rule. The values of the proportion classified in the group should be similar to the model estimates of group probabilities in the second column.

<sup>c</sup> Average posterior probability is obtained by averaging the posterior probabilities for a given group for all individuals placed in this group by the maximum posterior probability rule. Acceptable values for this criterion are 0.7 or greater for all groups.

<sup>d</sup> Acceptable values of the odds correct classification are 5.0 or greater for all groups.
Implications of “Trajectory Groups” & “Group Membership”

- Sample size and length of follow-up period influence the number of groups
- Goal: identify approximation of unique patterns, not the true number of groups
- Subjective decisions sometimes are necessary to identify a final trajectory model → transparency and disclosure of the decision are needed
- Group membership is a convenient statistical fiction, not a state of being
  - Individuals are not necessary following the exact group-level trajectory
  - Spaghetti plots can provide additional diagnostic information about model fitness and homogeneity by identifying through visual inspection
Other recommended steps in model selection

➢ Decide optimal order of groups for the “base specification” (e.g., all cubic, 1 linear and other cubics)
  o Use BIC if possible
  o Stop when the prominent features of data appear (by consulting clinicians or experts)

➢ Refine the trajectories order for the optimal number of group

➢ Require minimum % of the cohort assigned to each trajectory group (e.g., 1%, 2.5%, 5%) based on intervention needs

➢ When evaluating an association between trajectories and outcomes, minimum number of outcomes occurred in each trajectory may be required to stabilize the modeling (e.g., ≥2 cases)
At this point, you probably have many questions....

How do you select the number of groups/trajectories? How do you evaluate model adequacy?

How do you profile or describe group members?

Can you add time-invariant covariates to the trajectory itself?

Can you add time-varying covariates to the trajectory itself?

Can you describe two or more behaviors/outcomes at the same time?

What are differences between different methods to develop trajectory groups?
Profile Group Characteristics

➢ Conduct descriptive statistics by trajectory group

Table 1  Characteristics of Pennsylvania Medicaid enrollees with buprenorphine prescriptions and by trajectory group.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>All cohort (n = 10,945)</th>
<th>Discontinued at 1–3 months (n = 2,722)</th>
<th>Discontinued at 3–5 months (n = 2,053)</th>
<th>Discontinued at 5–8 months (n = 1,574)</th>
<th>Discontinued after 8 months (n = 1,420)</th>
<th>Refilled intermittently (n = 1039)</th>
<th>Refilled persistently (n = 2,337)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Socio-demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age, mean (SD)***</td>
<td>32.8 (9.3)</td>
<td>32.2 (9.2)</td>
<td>32.7 (9.4)</td>
<td>32.7 (9.4)</td>
<td>32.8 (9.2)</td>
<td>32.7 (9.2)</td>
<td>33.7 (9.1)</td>
</tr>
<tr>
<td>Female sex, n (%)</td>
<td>6379 (58.3)</td>
<td>1556 (57.2)</td>
<td>1218 (59.3)</td>
<td>789 (57.4)</td>
<td>813 (57.3)</td>
<td>629 (60.5)</td>
<td>1374 (58.8)</td>
</tr>
<tr>
<td>Race, n (%)***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>9784 (89.4)</td>
<td>2370 (87.1)</td>
<td>1809 (88.1)</td>
<td>1232 (89.7)</td>
<td>1279 (90.1)</td>
<td>953 (91.7)</td>
<td>2141 (91.6)</td>
</tr>
<tr>
<td>Black</td>
<td>508 (4.7)</td>
<td>155 (5.7)</td>
<td>128 (6.2)</td>
<td>55 (4.0)</td>
<td>56 (3.9)</td>
<td>34 (3.3)</td>
<td>80 (3.4)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>464 (4.2)</td>
<td>136 (5.1)</td>
<td>82 (4.0)</td>
<td>60 (4.4)</td>
<td>60 (4.2)</td>
<td>39 (3.8)</td>
<td>87 (3.7)</td>
</tr>
<tr>
<td>Others</td>
<td>189 (1.7)</td>
<td>61 (2.2)</td>
<td>34 (1.7)</td>
<td>27 (1.9)</td>
<td>25 (1.8)</td>
<td>13 (1.2)</td>
<td>29 (1.3)</td>
</tr>
<tr>
<td>Managed care health plan, n (%)***</td>
<td>7783 (71.1)</td>
<td>2075 (76.2)</td>
<td>1492 (72.7)</td>
<td>965 (70.2)</td>
<td>944 (66.5)</td>
<td>766 (73.7)</td>
<td>1541 (65.9)</td>
</tr>
<tr>
<td>Type of medical assistance eligibility, n (%)†</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GA</td>
<td>3248 (29.7)</td>
<td>846 (31.3)</td>
<td>618 (30.1)</td>
<td>424 (30.9)</td>
<td>441 (31.1)</td>
<td>301 (29.0)</td>
<td>618 (26.4)</td>
</tr>
<tr>
<td>SSI</td>
<td>3219 (29.4)</td>
<td>749 (27.5)</td>
<td>569 (27.7)</td>
<td>397 (28.9)</td>
<td>409 (28.8)</td>
<td>328 (31.6)</td>
<td>767 (32.8)</td>
</tr>
<tr>
<td>TANF</td>
<td>4229 (38.7)</td>
<td>1053 (38.7)</td>
<td>802 (39.1)</td>
<td>521 (37.9)</td>
<td>536 (37.7)</td>
<td>395 (38.0)</td>
<td>923 (39.5)</td>
</tr>
<tr>
<td>Waiver</td>
<td>249 (2.3)</td>
<td>74 (2.7)</td>
<td>64 (3.1)</td>
<td>32 (2.3)</td>
<td>34 (2.4)</td>
<td>15 (1.4)</td>
<td>30 (1.3)</td>
</tr>
<tr>
<td>Resided county, n (%)***</td>
<td>587 (5.4)</td>
<td>89 (3.3)</td>
<td>50 (2.4)</td>
<td>53 (3.9)</td>
<td>68 (4.8)</td>
<td>32 (3.1)</td>
<td>95 (4.1)</td>
</tr>
<tr>
<td>Rural</td>
<td>1539 (14.1)</td>
<td>345 (12.7)</td>
<td>277 (13.5)</td>
<td>199 (14.5)</td>
<td>227 (16.0)</td>
<td>124 (11.9)</td>
<td>367 (15.7)</td>
</tr>
<tr>
<td>Micropolitan</td>
<td>9019 (82.4)</td>
<td>2288 (84.1)</td>
<td>1726 (84.1)</td>
<td>1122 (81.7)</td>
<td>1125 (79.2)</td>
<td>883 (85.0)</td>
<td>1875 (80.2)</td>
</tr>
<tr>
<td>Metropolitan</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health-status</td>
<td>1.1 (1.4)</td>
<td>1.06 (1.38)</td>
<td>1.11 (1.38)</td>
<td>1.06 (1.41)</td>
<td>1.02 (1.31)</td>
<td>1.01 (1.3)</td>
<td>1.00 (1.28)</td>
</tr>
<tr>
<td>Elhbuahs comorbidity index (excl drug abuse diagnoses), range 0–30, mean (SD)</td>
<td>7371 (67.4)</td>
<td>1856 (68.2)</td>
<td>1397 (68.1)</td>
<td>927 (67.5)</td>
<td>953 (67.1)</td>
<td>701 (67.5)</td>
<td>1537 (65.8)</td>
</tr>
<tr>
<td>Opioid use disorder diagnosis, n (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Using multi-nominal logistic regression
At this point, you probably have many questions….

How do you select the number of groups/trajectories? How do you evaluate model adequacy?

How do you profile or describe group members?

- Can you add time-invariant covariates to the trajectory itself?
- Can you add time-varying covariates to the trajectory itself?
- Can you describe two or more behaviors/outcomes at the same time?
- What are differences between different methods to develop trajectory groups?
Statistically link group membership to individual characteristics

➢ Move beyond univariate contrasts
➢ Group identification is probabilistic, not certain
➢ Use of multinomial logit model to create a multivariate probabilistic linkage

\[ \pi_j(X_i) = \frac{e^{X_i\theta_j}}{\sum e^{X_i\theta_j}} \]

\( X_i \): covariate at baseline

Including time-invariant covariates in estimating group membership (Examples 6 & 7)

STATA Syntax

use http://www.andrew.cmu.edu/user/bjones/traj/data/montreal_sim.dta, clear
traj, var(qcp*op) indep(age*) model(cnorm) min(0) max(10) order(1 3 2)
[       risk(scolmer scolper)
trajplot, xtitle(Age) ytitle(Opposition)
]

Note: scolmer( mother’s schooling); scolper (father’s schooling)

2037 observations read.
203 excluded by if condition or by missing values in risk variables.
934 observations used in the trajectory model.

Maximum likelihood Estimates
Model: Censored Normal (cnorm)

| Group | Parameter | Estimate | Standard Error | T for H0: Parameter=0 | Prob > |T|
|-------|-----------|----------|----------------|-----------------------|--------|
| 1     | Intercept | 0.8460   | 0.35474        | 2.384                 | 0.0172 |
|       | Linear    | -0.17547 | 0.03865        | -5.745                | 0.0000 |
| 2     | Intercept | -11.09778| 4.27423        | -2.696                | 0.0040 |
|       | Linear    | 4.06666  | 1.36724        | 2.973                 | 0.0030 |
|       | Quadratic | -0.36582 | 0.13554        | -2.699                | 0.0070 |
|       | Cubic     | 0.00074  | 0.00026        | 2.286                 | 0.0230 |
| 3     | Intercept | -1.91449 | 1.02478        | -1.868                | 0.0610 |
|       | Linear    | 1.94556  | 0.20591        | 7.312                 | 0.0000 |
|       | Quadratic | -0.07512 | 0.00096        | -7.621                | 0.0000 |
|       | Sigma     | 2.58639  | 0.03404        | 75.987                | 0.0000 |

Group membership

1 Baseline (0.00000)
2 scolmer -0.04285 0.03898 -1.099 0.277
scolper -0.05367 0.03128 -1.716 0.0852
3 Constant 2.43032 0.46885 5.184 0.0000
scolmer -0.11564 0.04285 -2.685 0.0072
scolper -0.10886 0.03087 -4.225 0.0000

Log-odds estimates (can exponentiate it to get odds ratio)

Include covariates

No covariates

Effect of individual covariates on probability of trajectory group membership

- Covariates or risk factors on physical aggression: broken home at age 5, low IQ, low maternal education, mother began childbearing as a teenager

Graph allows nice contrast between no risk factors and all risk factors

At this point, you probably have many questions….

- How do you select the number of groups/trajectories? How do you evaluate model adequacy?
- How do you profile or describe group members?
- Can you add time-invariant covariates to the trajectory itself?
- Can you add time-varying covariates to the trajectory itself?
- Can you describe two or more behaviors/outcomes at the same time?
- What are differences between different methods to develop trajectory groups?
Including time-varying covariates in estimating group membership (Example 8)

STATA Syntax (the effect of gang membership on violent delinquency)

```
use http://www.andrew.cmu.edu/user/bjones/traj/data/gang_data_sim.dta, clear
matrix tc1 = 0, 0, 0, 0, 0, 0, 0
matrix tc2 = 0, 0, 0, 1, 1, 1, 1
traj, var(bat*) indep(*) model(zip) order(2 2 2 2) tcov(gang*) plotcov(tc1)
trajplot, xtitle(Scaled Age) ytitle(Rate)
traj, var(bat*) indep(*) model(zip) order(2 2 2 2) tcov(gang*) plotcov(tc2)
trajplot, xtitle(Scaled Age) ytitle(Rate)
```

**Predicted trajectories for not in a gang from age 11 to 17**

```
Predicted traitories for not in a gang from age 11 to 17
```

**Predicted trajectories for joining a gang starting at age 14**

```
Predicted trajectories for joining a gang starting at age 14
```

Adapted from Jones BL and Nagin DS. A Stata Plugin for Estimating Group-based Trajectory Models (https://sscindiana.edu/doc/wimdoc/2013-03-29_nagin_trajectory_stata-plugin-info.pdf)
Using Wald tests to examine differential time-varying factor effects by trajectory group (Example 8)

STATA Syntax (the effect of gang membership on violent delinquency)

/* List the parameter estimates to verify the names for testnl */
matrix list e(b), format(%8.3f)
testnl _b[gang89G1]=_b[gang89G5]
testnl _b[gang89G4]=_b[gang89G5]

. matrix list e(b), format(%8.3f)

e(b)[1,24]
    interc1  linear1  quadra1  gang89G1  interc2  linear2  quadra2  gang89G2  interc3  linear3  quadra3  gang89G3  interc4
   y2   5.549   -1.911    0.612   -5.469   10.433  -3.693    0.485    0.430   0.072  -0.157  -1.614

. testnl _b[gang89G1]=_b[gang89G5]
(1)  _b[gang89G1] = _b[gang89G5]
    chi2(1) =  37.22
    Prob > chi2 =  0.0000

. end of do-file
. do "C:\Users\wlociganic\AppData\Local\Temp\STD1b6c_00000.tmp"

. testnl _b[gang89G4]=_b[gang89G5]
(1)  _b[gang89G4] = _b[gang89G5]
    chi2(1) =  2.99
    Prob > chi2 =  0.0836

The coefficient estimates of gang effect differ for groups 1 and 5 (p<0.0001)

The coefficient estimates of gang effect differ for groups 1 and 4 (p=0.0836)

At this point, you probably have many questions....

- How do you select the number of groups/trajectories? How do you evaluate model adequacy?
- How do you profile or describe group members?
- Can you add time-invariant covariates to the trajectory itself?
- Can you add time-varying covariates to the trajectory itself?
- Can you describe two or more behaviors/outcomes at the same time?
- What are differences between different methods to develop trajectory groups?
Joint GBTM (Conditional Probability)

➢ Design to analyze the trajectory of two distinct but related outcomes

➢ To analyze connections between the developmental trajectories of two outcomes that are evolving contemporaneously (e.g., depression and alcohol use) or that evolve over different time periods (e.g., prosocial behavior in childhood and school achievement in adolescence)

➢ Key outputs:
  o Trajectory groups for both measurement series
  o The probability of membership in each identified trajectory
  o **Conditional probabilities** linking membership across trajectory groups of the two respective behaviors.

Stata: Joint GBTM (Example 13)

- Montreal data: the linkage of opposition behaviors from age 6 to 13 with property delinquency from ages 11 to 15.

### State: Joint GBTM (Example 13, continued)

**Group membership (model 1 group | model 2 group)**

<table>
<thead>
<tr>
<th>Group</th>
<th>(%)</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5.02072</td>
<td>1.86656 2.600 0.0072</td>
</tr>
<tr>
<td>2</td>
<td>90.46582</td>
<td>2.75608 32.802 0.00000</td>
</tr>
<tr>
<td>3</td>
<td>4.57346</td>
<td>1.43851 1.889 0.05890</td>
</tr>
<tr>
<td>4</td>
<td>0.00000</td>
<td>0.80089 0.011 0.9890</td>
</tr>
<tr>
<td>12</td>
<td>9.56663</td>
<td>2.35961 4.092 0.00000</td>
</tr>
<tr>
<td>22</td>
<td>62.11232</td>
<td>3.56239 17.614 0.00000</td>
</tr>
<tr>
<td>32</td>
<td>21.30089</td>
<td>3.15883 6.743 0.00000</td>
</tr>
<tr>
<td>42</td>
<td>6.39015</td>
<td>1.52323 4.550 0.00000</td>
</tr>
<tr>
<td>23</td>
<td>37.62210</td>
<td>4.74557 3.715 0.00002</td>
</tr>
<tr>
<td>33</td>
<td>46.68466</td>
<td>4.67178 9.993 0.00000</td>
</tr>
<tr>
<td>34</td>
<td>27.87968</td>
<td>5.30155 5.259 0.00000</td>
</tr>
<tr>
<td>43</td>
<td>7.81416</td>
<td>2.56922 3.041 0.00024</td>
</tr>
</tbody>
</table>

**Group membership (model 1 group | model 2 group)**

<table>
<thead>
<tr>
<th>Group</th>
<th>(%)</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16.2%</td>
<td>44.6%</td>
</tr>
<tr>
<td>2</td>
<td>14.6%</td>
<td>39.3%</td>
</tr>
<tr>
<td>3</td>
<td>42.7%</td>
<td>15.3%</td>
</tr>
<tr>
<td>4</td>
<td>42.1%</td>
<td>8.4%</td>
</tr>
<tr>
<td>12</td>
<td>56.1%</td>
<td>35.5%</td>
</tr>
<tr>
<td>13</td>
<td>5.0%</td>
<td>9.0%</td>
</tr>
<tr>
<td>24</td>
<td>64.8%</td>
<td>35.2%</td>
</tr>
</tbody>
</table>

**Group membership (model 1 group and model 2 group)**

<table>
<thead>
<tr>
<th>Group</th>
<th>(%)</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1%</td>
<td>4%</td>
</tr>
<tr>
<td>2</td>
<td>28.0%</td>
<td>28.5%</td>
</tr>
<tr>
<td>3</td>
<td>10%</td>
<td>10.3%</td>
</tr>
<tr>
<td>4</td>
<td>15%</td>
<td>1%</td>
</tr>
<tr>
<td>12</td>
<td>9.6%</td>
<td>9.8%</td>
</tr>
<tr>
<td>13</td>
<td>6%</td>
<td>6.2%</td>
</tr>
<tr>
<td>24</td>
<td>0%</td>
<td>3.2%</td>
</tr>
<tr>
<td>34</td>
<td>1.7%</td>
<td>1.7%</td>
</tr>
</tbody>
</table>

**Group membership (model 2 group)**

<table>
<thead>
<tr>
<th>Group</th>
<th>(%)</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9.9%</td>
<td>67.7%</td>
</tr>
<tr>
<td>2</td>
<td>17.4%</td>
<td>4.9%</td>
</tr>
<tr>
<td>4</td>
<td>4.9%</td>
<td></td>
</tr>
</tbody>
</table>

BIC=-13612.04 (N=8872)  BIC=-13575.89 (N=926)  AIC=-13498.59  ll=-13466.59

---

Adapted from Jones BL and Nagin DS. A Stata Plugin for Estimating Group-based Trajectory Models (https://ssrc.indiana.edu/doc/wim/docs/2013-03-29_nagin_trajectory_stata-plugin-info.pdf)
Stata: Multi-trajectory modeling (Example 14)

- Montreal data: opposition behaviors from age 6 to 13, property delinquency from ages 11 to 15, and aggression from age 6 to 16

STATA Syntax

```stata
use http://www.andrew.cmu.edu/user/bjones/traj/data/montreal_sim.dta, clear
traj, multigroups(3) var1(qcp84op-qcp91op) indep1(age1-age5) model1(cnorn) max1(10) order1(1 2 2) var2(qas'det)
 indep2(age3-age7) model2(zip) order2(2 1 2) var3(qcp*bat) indep3(age*) model3(cnorn) max3(6) order3(1 2 1)
multitrajplot, xtitle(Age) ytitle1(Opposition) ytitle2(Rate) ytitle3(Aggression) ylabel1(0(2)/6) ylabel2(0(1)/4) ylabel3(0(1)3)
```

**Group 1 (57.6%)**

- Group 1 (57.6) Group 2 (26.0%) Group 3 (16.5%)

---

At this point, you probably have many questions….

- How do you select the number of groups/trajectories? How do you evaluate model adequacy?
- How do you profile or describe group members?
- Can you add time-invariant covariates to the trajectory itself?
- Can you add time-varying covariates to the trajectory itself?
- Can you describe two or more behaviors/outcomes at the same time?
- What are differences between different methods to develop trajectory groups?
## Comparisons of Different Methods to Develop Trajectory Groups

<table>
<thead>
<tr>
<th>Methods</th>
<th>Descriptions</th>
</tr>
</thead>
</table>
| K-mean clustering | 1. Simpler and faster (less computational time)  
2. Longitudinal nature of the data is ignored |
| “Two-step” approach (i.e., mixed modeling + K-means clustering) | 1. Does not ignore the longitudinal nature of the data |
| Latent class analysis (LCA) | 1. Structural equation modeling (SEM)-based using latent variable (finite mixture modeling)  
2. Accommodate inter-individual variability (between-subjects) and intra-individual (within-subjects) patterns of change over time  
3. Assumption: data consist of ≥1 trajectory groups  
4. Have relative objective criteria (i.e., model fit indicators, e.g., BIC) |
| Latent class growth analysis (LCGA) | 1. Same as (1) to (4) in LCA  
2. Assumption: there is no within class variation (i.e., no random effects)  
3. GBTM is a LCGA (GBTM approximates an unknown distribution of individual differences with group)  
4. Software: SAS, STATA |
| Latent class growth mixture modeling (LCGMM) | 1. Same as (1) to (4) in LCA  
2. Assumption: there can be within class variation (including normally-distributed random effects; may cause computation difficulties)  
3. Usually has less groups identified than LCA  
4. Software: M-plus |

Comparisons of Different Methods to Develop Trajectory Groups

Generally, all classification methods revealed comparable trajectories

- GBTM, LCGA, and LCGMM seem to be preferable above the more simple methods (e.g., k-means clustering), all classification methods should be applied with great caution.
- The optimal solution for LCA and LCGA contained more classes compared with LCGMM.
- LCGMM may increase computational times.

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Features of popular software for longitudinal SEM (as at May 2013) (*Only those mentioned in Section 3 of this paper are reported).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software</td>
<td>SGE</td>
</tr>
<tr>
<td>Behaviou.r package/ procedure</td>
<td>Pisc. Traj</td>
</tr>
<tr>
<td>Model types</td>
<td>QSEM</td>
</tr>
<tr>
<td>Outcome types and link function</td>
<td>Censored normal</td>
</tr>
<tr>
<td>Continuous (censored and nominal)</td>
<td>#</td>
</tr>
<tr>
<td>Binary</td>
<td>Logit</td>
</tr>
<tr>
<td>Grant</td>
<td>Poisson, Zero inflated</td>
</tr>
<tr>
<td>Trajectory specification</td>
<td>Censored normal only</td>
</tr>
<tr>
<td>Random effects</td>
<td>GLLAMM</td>
</tr>
<tr>
<td>Covariance structure of random effects</td>
<td>Censored normal, mixed between classes</td>
</tr>
<tr>
<td>(1 matrix)</td>
<td>GLLAMM: covariates may be specified by user</td>
</tr>
<tr>
<td>R matrix</td>
<td>Fixed to be the same across classes and time</td>
</tr>
<tr>
<td>Allows for first-order cross-sectional variance in R</td>
<td>GLLAMM</td>
</tr>
<tr>
<td>Fit criteria and test statistics</td>
<td>AIC, AICc, BIC, log-likelihood, Wald test</td>
</tr>
<tr>
<td>Fit and test statistics*</td>
<td>GMLAMM</td>
</tr>
</tbody>
</table>

At this point, you probably have many questions…. 

<table>
<thead>
<tr>
<th>Question</th>
<th>Answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>How do you select the number of groups/trajectories? How do you evaluate model adequacy?</td>
<td>How do you profile or describe group members?</td>
</tr>
<tr>
<td>How do you profile or describe group members?</td>
<td>Can you add time-invariant covariates to the trajectory itself?</td>
</tr>
<tr>
<td>Can you add time-invariant covariates to the trajectory itself?</td>
<td>Can you add time-varying covariates to the trajectory itself?</td>
</tr>
<tr>
<td>Can you add time-varying covariates to the trajectory itself?</td>
<td>Can you describe two or more behaviors/outcomes at the same time?</td>
</tr>
</tbody>
</table>
Summary of GBTM

➢ Identify and visualize groups following similar dynamic changes in medication utilization or other measures over time
➢ Transparency and disclosure of the decision for final model section are needed
➢ Equivalent or better prediction performance
➢ Different trajectories may have different characteristic profiles
➢ May better inform and guide target interventions and clinical management
Acknowledgement

➢ Special thanks to my biostatistics advisor Dr. Roslyn Stone

➢ Research Teams and Collaborators
  - **University of Florida:** Haesuk Park, Almut Winterstein, Juan Hincapie-Castillo, Serena Guo, James Huang, Gary, Resfield, Patrick Tighe, Chris Harle, Ron Shorr, Jiang Bian, Yonghui Wu, Khoa Nguyen, Debbie Wilson, Peggy Chang, and Seonkyeong Yang
  - **University of Pittsburgh:** Walid Gellad, Julie Donohue, Courtney Kuza, & Qingnan Yang
  - **Carnegie Mellon University:** Jeremy Weiss & Rayid Ghani
  - **University of Arizona:** Kent Kwoh, Jeannie Lee, Helen Zhang, & Ge Yong
  - **University of Utah:** Adam Gordon, Jerry Cochran & Dan Malone
  - **University of Kentucky:** Chris Delcher
Thank you!

Jenny Lo-Ciganic MSPharm, MS, PhD, wlociganic@cop.ufl.edu