GAME OF PHONES: A SONG OF DAILY REPORTS WITH MISSING OBSERVATIONS AND THE MODELS THAT ANALYZE THEM

W. Scott Comulada
UCLA Dept of Psychiatry & Biobehavioral Sciences
wcomulada@mednet.ucla.edu

November 18, 2014

CHIPTS Methods Seminar

Cell phone-based data collection

WHY?

- Cell phones are already part of our daily routine
- Easy for individuals to report in the moment on a daily basis
  - Referred to as ecological momentary assessment (EMA)
  - Will refer to cell phone-based EMA as CEMA
- Event-oriented recording
  - Drug use and other episodic behaviors (Shiffman, 2009)
- Consider the alternative: retrospective self-report
  - Interviewers in a laboratory setting
  - Recall biases (Bradburn, Rips, & Shevell, 1987; Stone & Shiffman, 1994)

Cell phone-based data collection

WHY?

Blends nicely with technology-based interventions

- Ecological momentary intervention (Heron & Smyth, 2010)
- Key aspects:
  - Better self-awareness of problem behaviors due to frequent assessment
  - Ideally, self-monitoring leads to behavior change (reactivity)
  - Often a small effect in practice
    - EMA of undergraduate problem drinkers (Hufford et al., 2002)

Cell phone-based data collection

Caveat: What is acceptable participant burden?

- Retrospective recall burden:
  - Scheduling clinic visits (not with web-based assessment)
  - Commitment over longer period of time (often several years) relative to (C)EMA
- EMA burden: Daily assessment (often 3, 4x a day)
- As with any data collection method, know your population

Cell phone-based data collection

Candidate populations for CEMA

- Based on our own experience and compliance rates

Patient populations (compliance rates of 80% and better)

  - Crack-addicted homeless patients (Freedman et al., 2006)
  - Youth in outpatient drug treatment (Comulada et al., accepted)

- versus non-patient populations (closer to 50%)
  - HIV-positive adults recruited in L.A. (Swendeman et al., accepted)
  - Australian youth recruited from a health clinic (Kauer et al., 2009)

We can only minimize CEMA burden so much, and so an UNWELCOME GUEST of LONGITUDINAL data collection remains…

MISSING DATA!

A brief overview of missing data issues

Mechanisms for missing data
(Rubin, 1976)

Ignorable

Non-

Missing Completely at Random (MCAR)

Missing at Random (MAR)
[Dependent on observed data]

Missing Not at Random (MNAR)
[Dependent on unobserved data]

Consequences of missing data

MCAR

Smaller sample size (less power) upon which to make inferences

MAR and MNAR

Biased estimation

- Draw incorrect conclusions from data
- For example, evaluation of a substance use intervention will be impacted by non-response if users do not fill out CEMA when using
- See McPherson et al. (2012) for further discussion

MAR data tools for longitudinal data

Multiple imputation

Regression model

- Working correlation matrix: Generalized Estimating Equations (GEE)
  - e.g., PROC GENMOD in SAS
  - Adjusts for missing covariate observations only
- Random effects models: (L)ikelihood and (B)ayesian methods
  - e.g., L: PROC MIXED; B: WinBUGS, JAGS
  - Adjusts for missing covariate & outcome observations

NMAR data tools for longitudinal data

Two main approaches

1. Pattern mixture models (Little, 1993, 1994, 1995; Hedeker & Gibbons, 1997)

Include missing data patterns as covariates (or latent classes; Muthen et al., 2011)

- We have a longitudinal data set with 3 time points
- The following patterns of observed (O) and missing data (M) occur

Time 1	Time 2	Time 3
O	O	O
O	O	M
O	M	M
O	M	O

We fit a regression model with 3 dummy variables (O O O is reference group)

CEMA data: ↑ # of time points → ↑↑↑ # of possible patterns

- Pattern mixture models less practical
- Today we focus on second NMAR modeling approach

NMAR data tools

Two main approaches

2. Selection models (e.g., Diggle & Kenward, 1994; Wu & Carroll, 1988)

Idea: Fit longitudinal Bivariate-outcome model

- Outcome for observed data as you normally would (random effects account for MAR)
- Outcome for presence or absence of missing data (Y/N)
- Covariance structure to model correlations between both outcomes

Data structure: Binary outcome; 1 participant; 3 possible time points →

Two outcomes	Time 1	Time 2	Time 3
Obs. data	1	.	0
Missing indicator	0	1	0

Missing data already discussed a lot in the literature

Why revisit in the context of CEMA?

How do we define CEMA compliance?

- Not as easy as in traditional study: Do not show up for interview → non-compliant
- e.g., CEMA study of youth in outpatient drug treatment (Comulada et al., accepted)
- For each CEMA question, prompted several times if no response
- Most complete assessment in a few minutes, but some took hours
- Do we discard assessments that took more time?

Different missing data patterns that are problematic

- Traditional study over several years: Drop out
- CEMA: May be little drop out (in treatment program, shorter time frame)
- Intermittent missing data (not filling out CEMA during study)

Missing data already discussed a lot in the literature

Why revisit in the context of CEMA?

Different models are appropriate (e.g., covariance structures)

- Traditional: Few time points → Random-intercept
- CEMA: Intensive longitudinal data (ILD; Walls & Schafer, 2006) → Autoregressive structures

Longitudinal models for NMAR data in general (traditional and CEMA) have not adequately been addressed

Consider limitations of commercial software to fit NMAR models:

Continuous outcomes

- "Normally" distributed psychological scales in MPLUS (Muthen et al., 2011)
- Yet public health outcomes are often discrete,

  binary indicator for drug use, count of unprotected sex acts
- Debatable how many psychological scales are normally distributed (Cliff and Keats, 2003)

Single outcomes

- "Chicken and egg problem" (Comulada, Muth, & Latkin, 2012). Specification of predictor-outcome assumes relationship is already understood
- Often preferable to model multiple outcomes
- e.g., self-report (CEMA and/or retrospective report) and urine screenings are conducted during substance abuse treatment
- Which one should be treated as outcome / predictor?

Software implementation issue:

Based on maximum likelihood (ML) methods

- Suffices for many standard regression analyses
- Convergence issues for complex models such as longitudinal NMAR models

Solution: Bayesian methods (FOCUS of today's seminar)

Of course, not a perfect solution

- Computational burden still present
  - Shifted from maximization to simulation
  - Decreasing barrier as computational speeds increase
- Lacks convergence criteria
  - Diagnostic plots are key
- Specification of prior distribution (+ specification of outcome distribution that is done for ML methods)
  - Sensitivity of results to prior specification is needed

Illustration using Bayesian software to fit longitudinal NMAR model

Motivating study data

24 youth in outpatient drug treatment program

- Pilot tested Text message-based / once-a-day CEMA
- Queried on alcohol and other drug use (AOD), sexual behavior, and context (e.g. feelings during use).
- Assigned to receive “daily” end-of-day (at 9PM) or random (1x between 3 to 9PM) CEMA

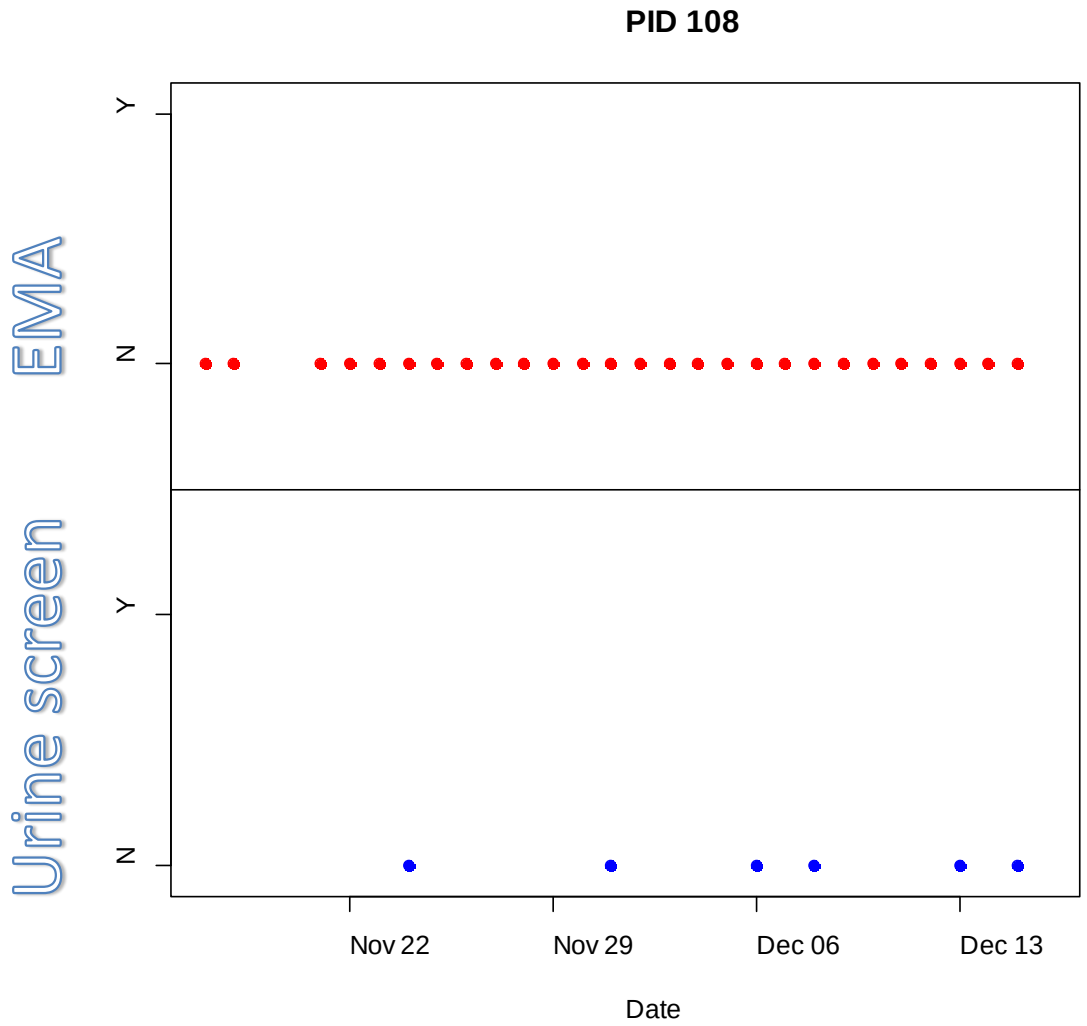
Motivating study data

Data structure

- Eligibility required participants to be in treatment for at least a month (for one rotation)
- During rotation: Daily EMA, urine screenings (2x week)
- Could participate in up to 4 rotations (month rest in between)
- Incentives: $25 a week / 500 free phone minutes a month
- High interest to participate: Free minutes, coolness factor of phone

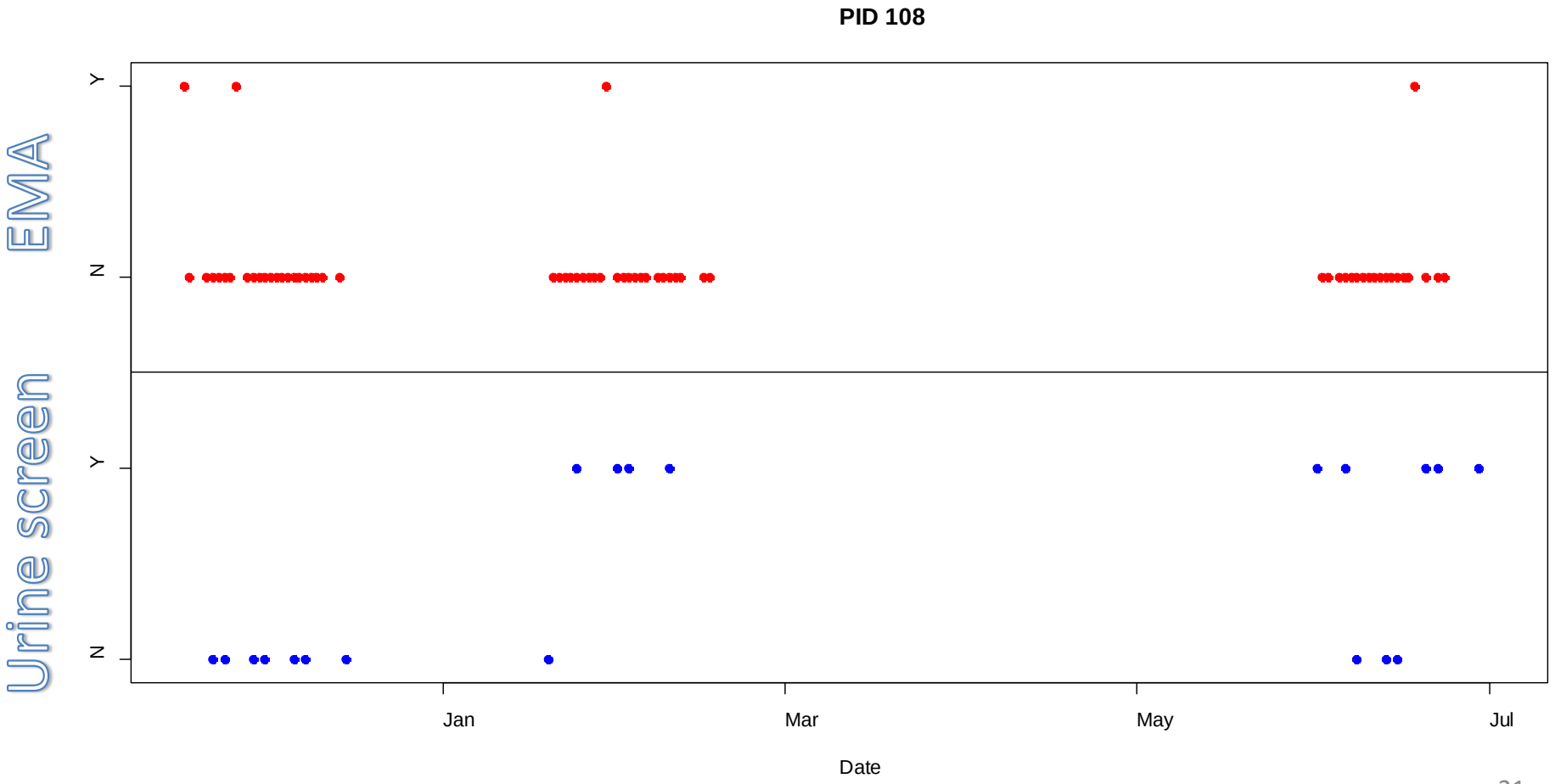
Motivating study data

Example of reported **marijuana** use by participant in study for 1 rotation



Motivating study data

Example of reported **marijuana** use by participant in study for multiple rotations



Motivating study data

Median and maximum number of days where AOD use was reported out of daily EMA reports (N = 24 participants)

	Median		Maximum	
Substance	**n**	**%**	**n**	**%**
Alcohol	3	6.7%	18	62.1%
Marijuana	2	5%	22	58.6%
Ecstasy	0	0%	2	5%
Cocaine	0	0%	2	3.5%
Inhalants	0	0%	13	22.8%
LSD	0	0%	1	1.8%
Painkillers	0	0%	3	5%
Meth	0	0%	2	6.7%

Motivating study data

Median and maximum number of days where positive urine screens occurred out of urine screens (N = 24 participants)

	Median		Maximum	
Substance	**n**	**%**	**n**	**%**
Alcohol	0	0%	6	83%
Marijuana	1.5	10%	14	100%
Cocaine	0	0	2	40%
Meth	0	0	6	60%
Any drug	1.5	12%	15	100%

Motivating study data

- 12% of EMA reports were missed out of 1178 possible reporting days
- 18% of urine screens were missed out of 310 possible bi-weekly screens
- No drop-out during rotation: Captive audience (In treatment for at least one month; only intermittent missing data)

<u>Overall</u>

- ↓ Drug use (expected from youth were in treatment)
- ↑ Compliance

Analysis

Model marijuana use

- Indicated by self-report (CEMA)
- Indicated by positive urine screening

Model non-response for both outcomes as well

Bayesian modeling software

Models fit in **JAGS** (Just Another Gibbs Sampler)

http://mcmc-jags.sourceforge.net/)

- Similar to WinBUGS / OpenBUGS
- Downside: Command-line interface (the ugly black dialogue box)
- Solution: Run JAGS through **R** using library: **rjags**

(http://cran.r-project.org/web/packages/rjags/index.html)

Install JAGS, install rjags in R, type commands in R

Selection model notation

Binary outcome y_{itk} for a person in rotation *i* at time *t* for outcome k:
EMA reported use (k = 1), EMA non-response (k = 2), + urine screening (k = 3), and screening non-response (k = 4)

Candidate logistic models for $p_{itk} = p(y_{itk} = 1)$

(a) Separate random effects (RE) for each outcome

$$\text{logit}(p_{itk}) = X'_{itk}\beta_{tk} + \eta_{ik}$$

Note: Small values for covariance parameters in V, except between outcomes for EMA reported use (k = 1) & screening (k = 3)

(b) Shared RE for k = 1, 3 that differs by scale parameter λ

$$\text{logit}(p_{it1}) = X'_{it1}\beta_{t1} + \eta_{i1}\ ,\ \ \text{logit}(p_{it3}) = X'_{it3}\beta_{t3} + \lambda\eta_{i1}$$

$$\text{logit}(p_{it2}) = X'_{it2}\beta_{t2} + \lambda\eta_{i2}\ ,\ \text{logit}(p_{it4}) = X'_{it4}\beta_{t4} + \lambda\eta_{i4}$$

Selection model covariates

- Time in rotation
- Rotation type (end-of-day vs. random)
  - Rotation type not included in screening part of model
- EMA use outcome as covariate for EMA non-response outcome
- Urine screening outcome as covariate for screening non-response outcome
  - Non-significant regression coefficient on outcome covariate indicates lack of evidence for NMAR data
- Age and ethnicity were not included (little variation); no gender differences were found

Selection model (a): Snippets of JAGS code

```
for( i in 1:1178 ){

 emaruse[i] ~ dbern(up01[i])
 logit(up01[i]) <- inprod(uX[i,],ua) + re[upidr[i],1]

 emarmiss[i] ~ dbern(upna[i])
 logit(upna[i]) <- inprod(uXm[i,],ub) + emaruse[i]*uc + re[upidr[i],2]}
```

Included to test NMAR

Modeling 1178 EMA responses

```
for( i in 1:310 ){

 smaruse[i] ~ dbern(sp01[i])
 logit(sp01[i]) <- sX[i]*sa + re[spidr[i],3]

 smarmiss[i] ~ dbern(spna[i])
 logit(spna[i]) <- sXm[i]*sb + smaruse[i]*sc + re[spidr[i],4] }
```

Included to test NMAR

Modeling 310 urine screens

```
for( k in 1:41 ){
re[k,1:4]    ~ dmnorm(mu[1:4],V[1:4,1:4]) }
```

Random effects between EMA / screens

Diagnostic plots for Bayesian analyses

Basic idea: Model parameters are described through posterior distribution based on prior distribution & likelihood

Simple case: Solution is readily available

- Binary variable and events occur with probability θ
- If we specify prior for θ to be beta distribution then posterior distribution for θ is also beta distribution

Complex models: Use Markov Chain Monte Carlo (MCMC) techniques to mimic random samples from posterior distribution

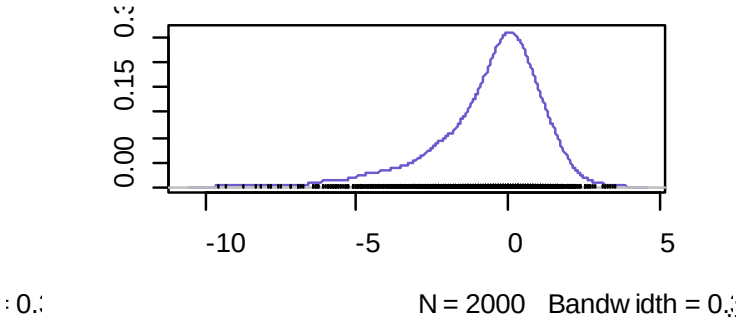
- Calculate means of random samples as parameter estimates (and other summary statistics)
- Diagnostic plots needed to ensure that samples are reasonably random (not too corrrelated) and from posterior distribution
- Quite a few diagnostic tests and plots that are available
- Here we show a few basic ones

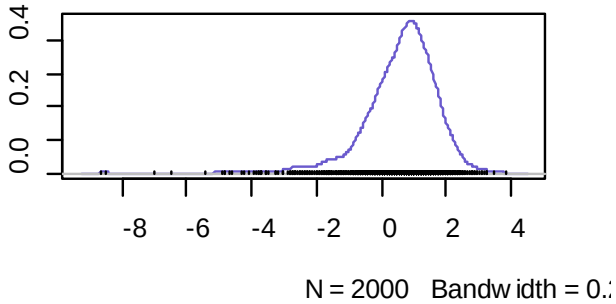
Diagnostic plots for EMA drug use covariate

Model (a) Separate RE

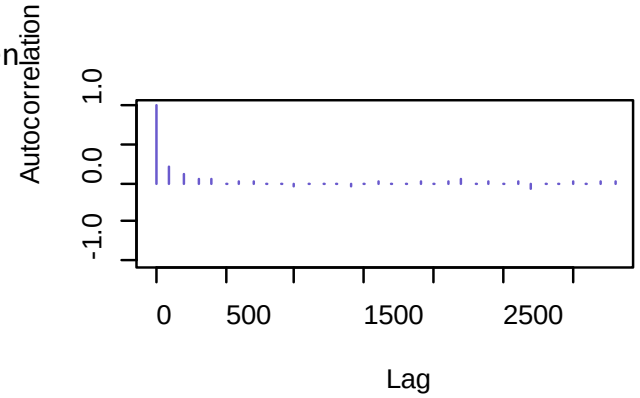
Model (b) Shared RE

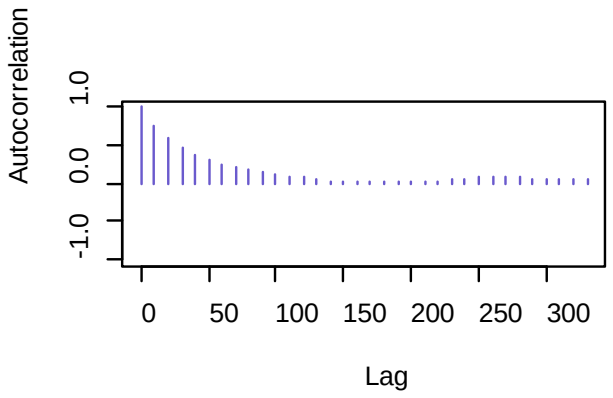
Density plots
Ideal: Normal dist.



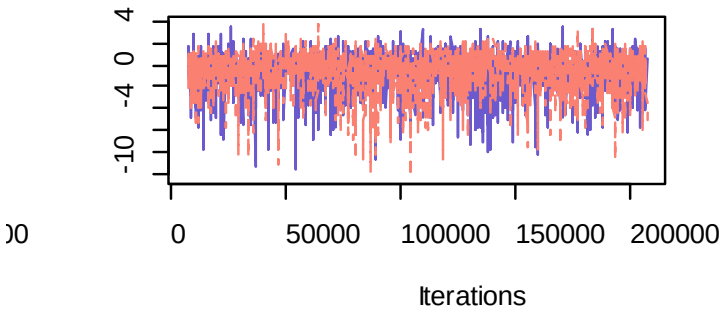


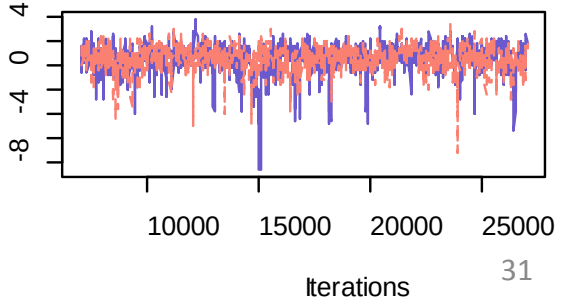
Autocorrelation plots
Ideal: Correlation of samples between lags drops to zero quickly





Trace plots of 2 chains
Ideal: a lot of overlap





Selection model (b) results for marijuana use (Y/N) Means (Estimate) and standard deviations (SD) of random samples from posterior distribution

Parameters		Est.	SD
Use (Y/N)	Intercept	-3.35	0.44
	Daily vs. Random	-1.19	0.44
+ Screen (Y/N)	Intercept	-2.92	0.83
Random effects	Variance	4.00	1.59
	Scale (+ Screen)	1.94	0.45
Non-resp. (Use)	Intercept	-3.50	0.36
	Days in rotation	0.062	0.012
	Use (Y/N)	0.57	1.08
Non-resp. (Scr)	Intercept	-1.65	0.30
	+ Screen (Y/N)	-0.70	1.57
Random effects	Var(Non-r Use)	1.46	0.51
	Var(Non-r Scr)	0.93	0.53
	Correlation	0.31	0.24

Selection model results for marijuana use (Y/N)

Parameters		Est.	SD
Use (Y/N)	Intercept	-3.35	0.44
	Daily vs. Random	-1.19	0.44
+ Screen (Y/N)	Intercept	-2.92	0.83
Random effects	Variance	4.00	1.59
	Scale (Screen)	1.94	0.45
Non-resp. (Use)	Intercept	-3.50	0.36
	Days in rotation	0.062	0.012
	Use (Y/N)	0.57	1.08
Non-resp. (Scr)	Intercept	-1.65	0.30
	+ Screen (Y/N)	-0.70	1.57
Random effects	Var(Non-r Use)	1.46	0.51
	Var(Non-r Scr)	0.93	0.53
	Correlation	0.31	0.24

↓ Use ↔ Daily rotation

Likely result of way use was queried ("Today" for daily; "Since last survey" for random)

↑ Non-resp. ↔ time in rotation

Not surprising (fatigue)

Don't see evidence of NMAR

Conclusions from results

Don’t see evidence of NMAR based on this model

i.e., if missingness at given time point depends on use at same time point

Missingness may depend on use at other time points

Possibly difficult to detect in this sample

High compliance and low drug use

Correlations between random effects

High between use / + Screen

Speaks to agreement between different measures of use

Low between non-response for EMA / missing a screen

Different reasons for missing

e.g., screen requires an office visit whereas EMA does not

References

Bradburn, .NM., Rips, L.J., & Shevell, S.K. (1987). Answering autobiographical questions: The impact of memory and inference on surveys. *Science*, 236(4798), 157-161.

Cliff, N., & Keats, J.A. (2003). Ordinal measurement in the behavioral sciences. Lawrence Erlbaum Associates, Mahwah, NJ.

Comulada WS, Muth SQ, & Latkin CA. (2012). The analysis of multiple ties in longitudinal egocentric network data: A case study on bidirectional relationships between trust and drug use. Social Networks, 34, 691-700.

Comulada WS, Lightfoot M, Swendeman D, Grella C, & Wu N. Compliance to cell phone-based EMA among Latino youth in outpatient treatment. Journal of Ethnicity in Substance Abuse. Accepted.

Diggle, P., Kenward, M.G. (1994). Informative drop-out in longitudinal data analysis. *Applied Statistics-Journal of the Royal Statistical Society Series C*, 43, 49-93.

Freedman, M.J., Lester, K.M., McNamara, C., Milby, J.B., & Schumacher, J. E. (2006). Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. *Journal of Substance Abuse Treatment*, 30(2), 105-111.

Hedeker, D., Gibbons, R.D. (1997). Application of random-effects pattern-mixture models for missing data in longitudinal studies. *Psychological Methods*, 2, 64-78.

Heron, K.E., Smyth, J.M. (2010). Ecological momentary interventions: incorporating mobile technology into psychosocial and health behavior treatments. *Br J Health Psychol*, 15, 1-39.

Kauer, S.D., Reid, S.C., Sanci, L., & Patton, G.C. (2009). Investigating the utility of mobile phones for collecting data about adolescent use and related mood, stress and coping behaviors: Lessons and recommendations. *Drug and Alcohol Review*, 28(1), 25-30.

Hufford, M.R., Shields, A.L., Shiffman, S., Paty, J., & Balabanis, M. (2002). Reactivity to ecological momentary assessment: An example using undergraduate problem drinkers. *Psychology of Addictive Behaviors*, 16(3), 205-211.

References

Little, R.J.A. (1993). Pattern-mixture models for multivariate incomplete data. JASA, 88, 125-133.

Little, R.J.A. (1994). A class of pattern-mixture mdoels for normal incomplete data. Biometrika, 81, 471-483.

Little, R.J.A. (1995). Modeling the drop-out mechanism in repeated-measures studies. JASA, 90, 1112-1121.

McPherson, S., Barbosa-Leiker, C., Burns, G.L., Howell, D., Roll, J. (2012). Missing data in substance abuse treatment research: Current methods and modern approaches. *Exp Clin Psychopharmacol*, 20, 243-250.

Muthen, B., Asparouhov, T., Hunter, A.M., Leuchter, A.F. (2011). Growth modeling with nonignorable dropout: Alternative analyses of the STAR*D antidepressant trial. *Psychol Methods*, 16, 17-33.

Rubin, D.B. (1976). Inference and missing data. Biometrika; 63: 581-592.

Shiffman, S. (2009). Ecological momentary assessment (EMA) in studies of substance use. Psychological assessment; 21: 486-497.

Swendeman D, Comulada WS, Ramanathan N, Lazar M, & Estrin D. Reliability and validity of daily self-monitoring by smartphone application for health-related quality of life, antiretroviral adherence, substance use, and sexual behaviors among people living with HIV (2014). AIDS and Behavior. Epub ahead of print.

Walls, T.A., Schafer, J.L. (2006). *Models for Intensive Longitudinal Data*. Oxford University Press.

Wu, M.C., Carroll, R.J. (1988). Estimation and comparison of changes in the presence of informative right censoring by modeling the censoring process. *Biometrics*, 44, 175-188.

Thank you

W. Scott Comulada

wcomulada@mednet.ucla.edu