Scaled Inverse Probability Weighting for Event Non-Reporting in Ecological Momentary Assessment Studies

Stephanie A. Kovalchik,
skovalch@rand.org

CHIPTS Methods Seminar

July 13, 2015

RAND Collaborators

Steven Martino

Rebecca Collins

Kirsten Becker

Elizabeth DAmico

William Shadel

Marc Elliott

This research was supported by NIAAA Grant R01AA021287

Outline

Definition

Motivation

Problem of Non-reporting

Correcting for Non-reporting

Application

Summary

Definition

> *The hallmark of ecological momentary assessment (EMA)—also known as event or experience sampling—is the collection of repeated momentary assessments from participants in their natural environments.*

—Shiffman & Stone (1998). *Health Psychology*, 17(1).

How is EMA Implemented?

Today, EMA data collection is most often conducted with the use of a handheld computing device.

Traditional Assessment

EMA Assessment

Advantages of EMA

Advantages of EMA

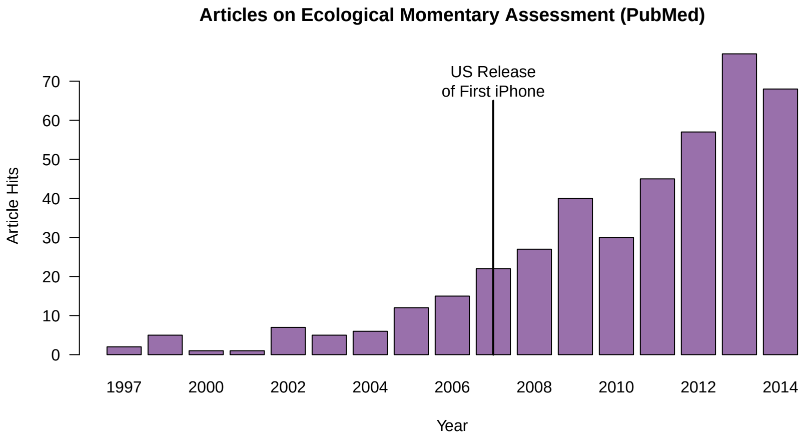
In contrast to traditional recall-based data collection,

1. Assessments occur at the time of the focal event
2. Multiple assessments are collected from each subject
3. Data collection occurs in the natural environment

These features help to

- Reduce recall bias
- Control for individual variation
- Reduce ‘white coat’ bias

Growing Popularity of EMA Designs



Common Applications

- Smoking cessation
- Alcohol consumption
- Insomnia
- Assessment of chronic pain
- Eating disorders

Adolescent Exposure to Alcohol Advertisement

- The motivation for the present work comes from the Tracking and Recording Alcohol Communications Study (TRAC).
- The TRAC study used an EMA design to investigate momentary shifts in youths alcohol-related attitudes and cognitions that may occur in response to real-world exposures to alcohol advertisements at the time of exposure.

Adolescent Exposure to Alcohol Advertisement

Study Objectives

- Research based on Nielsen ratings suggests that youth are exposed to 1 television ad for alcohol per day, on average.
- The objective of the present study was to expand estimates to include all media types,
- and to use EMA to obtain more accurate estimates of exposure.

Overview of TRAC Study Design

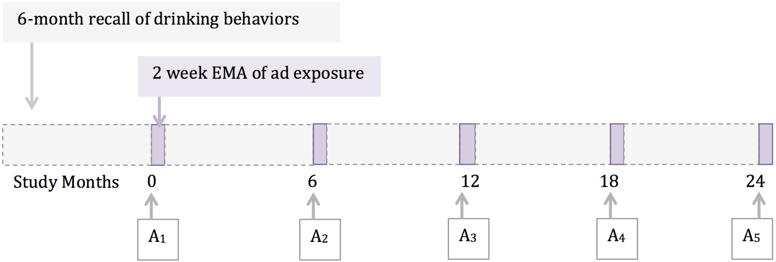
Sample

- 700 6th to 8th graders
- 3 cohorts with 3-month staggered recruitment
- Ethnically diverse
- Recruited from Los Angeles, CA School Districts

Data Collection

- 2-year longitudinal study
- Data collection at 6-month intervals
- Two types of surveys:
  1. Paper survey of behaviors
  2. Smartphone EMA of ad exposure

EMA Design


6-month recall of drinking behaviors
2 week EMA of ad exposure
Study Months
0
6
12
18
24
A_1
A_2
A_3
A_4
A_5

Assessment Instructions

- Participants underwent a one-day training session on the operation of the handheld devices (e.g. Samsung Galaxy Player 3.6)

- Participants were instructed to:
  1. Keep their device turned on at all times
  2. Charge the device at night
  3. Initiate data entry each time they encounter an alcohol ad
  4. Respond to random prompts

Challenges with EMA

- Lack of compliance
- Skipped event reports
- Over-Reporting

⋆ Greater concerns with special populations, like adolescents or addicts

Non-Compliance vs. Non-Reporting

Compliance

- *Compliance* refers to the completion of a scheduled random prompt.
- Non-compliance occurs when a participant fails to complete a scheduled assessment.
- Missed scheduled assessments are known to the researcher.
- Non-compliance reduces the number of available control assessments.

Non-Compliance vs. Non-Reporting

Non-Reporting

- *Non-reporting* is the failure to report an alcohol ad.
- Non-reporting is a type of ‘hidden’ missingness, because the timing of ads are not known to the researcher.
- Non-reporting reduces the number of event reports, which could bias estimates of event rates and associations.

Strategies to Prevent Non-Response

- Reminders
- Incentives for compliance
- Shorten surveys to reduce response burden

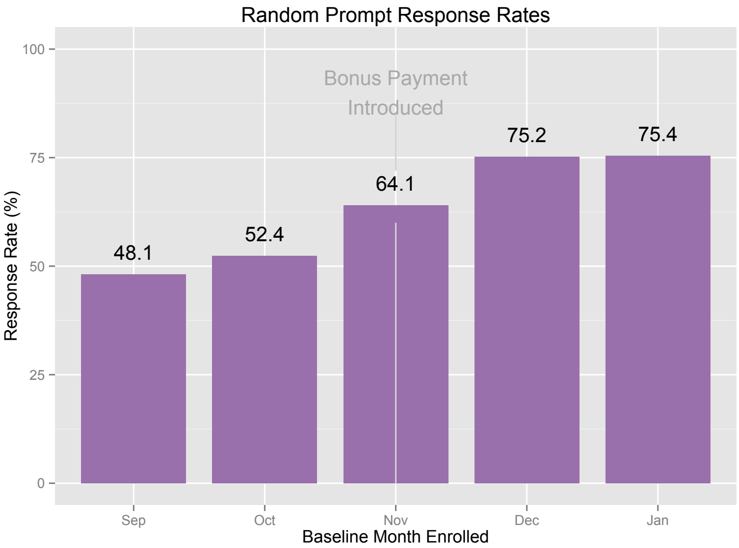
Randomized Short Forms

Domain	Question	Full Version	Short Version I	Short Version II	Short Version III
Exposure	Q1				
	Q2				
	Q3				
	Q4				
	Q5				
	Q6				
	Q7				
	Q8				
	Q9				
Perceptions	Q1				
	Q2				
	Q3				
Willingness	Q1				
	Q2				
	Q3				
Norms	Q1				
	Q2				
	Q3				
Expectations	Q1				
	Q2				
	Q3				
	Q4				
	Q5				
	Q6				

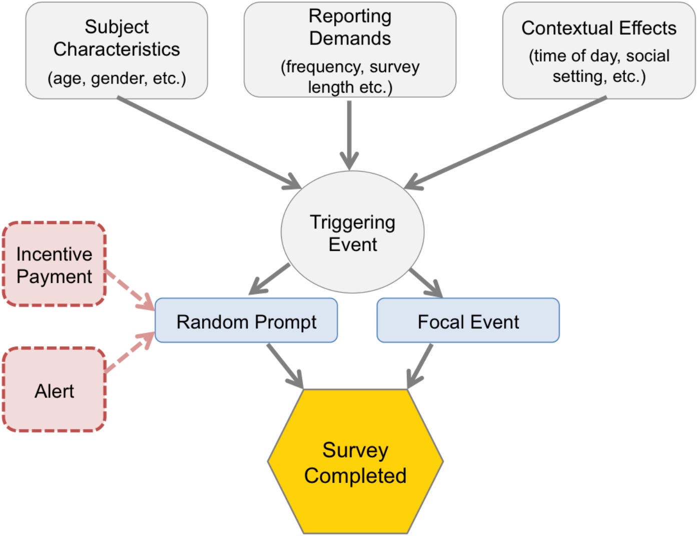
Randomization Scheme

	Estimated Quartile of Ad Exposure			
	Q1	Q2	Q3	Q4
Full survey	0.400	0.300	0.200	0.100
Short-form survey I	0.240	0.280	0.320	0.360
Short-form survey II	0.210	0.245	0.280	0.315
Short-form survey III	0.150	0.175	0.200	0.225

Effect of Incentive Payment



Conceptual Model for Reporting



Handling Non-Response

Since prevention strategies are unlikely to eliminate missed reports, we consider strategies to correct for report non-response analytically. There are three main components to our proposed strategy:

- **Monitoring** with scheduled random assessments
- **Investigating** reasons for non-response
- **Adjusting** with non-response weights

Observed Response Patterns

Our conceptual model supposes that compliance and event reporting patterns might differ.

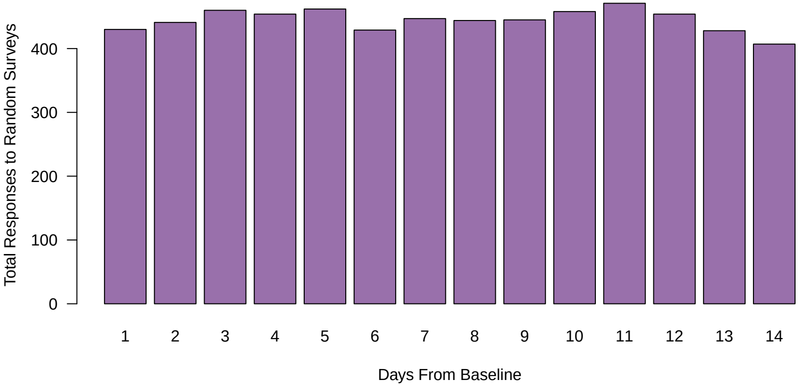
- Do we observe a difference in the TRAC study?
- What are the compliance and reporting patterns we observe?

Compliance Pattern: Subject Illustration

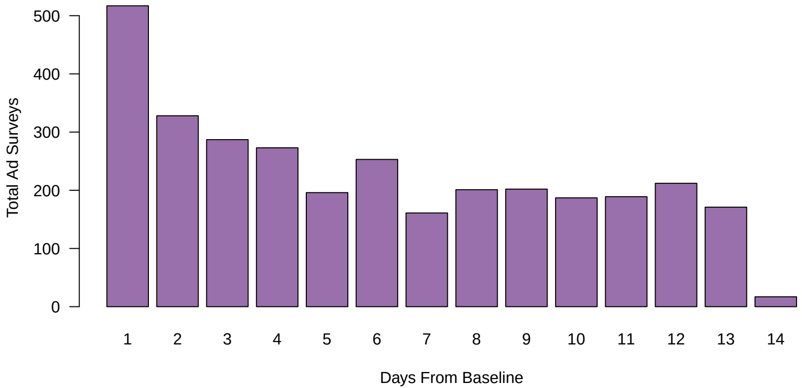
Random Assessment Completes: Subject 1002

Period	Study Day													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Morning														
Afternoon														
Evening														

Overall Compliance with Random Prompts



Ad Reporting



Remarks

- Completion of random prompts was steady throughout the study period
- Ad reports decline sharply with study day, suggesting strong fatigue effects
- The attrition pattern in total reports means that the observed ad events are incomplete and reporting is more incomplete later in the study than earlier in the study

Correcting for Non-Reporting

- The observed reporting patterns suggest that an ad event occurring at time t has some probability, $\pi(t)$, of being reported
- If $\pi(t)$ were known, the inverse probability weight $w(t) = 1/\pi(t)$ could be used to correct for missing reports
- We therefore sought an approach to estimate $\pi(t)$

Notation

Subjects $i = 1, 2, \ldots$

Report times $t_1, \ldots, t_{m_i}$

Study days $d = 1, 2, \ldots$

Subject and contextual characteristics $\mathbf{X}_i(t)$

Response indicators to random prompts $\mathbf{A} = (a_{i1}, \ldots, a_{ip})$

Two-Component Reporting Model

We model the reporting weight for the report at time t as the product of a compliance weight and scaling factor,

$$w_i(t) = \alpha(\mathbf{X}_{1i}(t))\beta(\mathbf{X}_{2i}(t)) \tag{1}$$

$$\psi(\alpha) = 0 \text{ and } \alpha(t) > 0 \ \forall t \tag{2}$$

- β is a compliance weight [Component 1]
- α is a positive scaling factor to adjust for fatigue [Component 2] effects
- $\psi(\alpha)$ is a set of constraints to identify the scaling factor α

A two-step process obtains estimates for each component to obtain an estimate for the final response weight.

Compliance Weight

The compliance weight is the inverse probability of response to the random prompt at time t,

$$P(a(t)|\mathbf{X}_{2i}(t))^{-1} = \beta(\mathbf{X}_{2i}(t)). \tag{3}$$

- Regression models for binary repeated measures would be appropriate for estimating this probability from the observed $\mathbf{A}(t)$.
- β would be estimated as the inverse of the predicted compliance probabilities.

Scaling Factor (1)

- Because we lack information about missing event reports, we instead specify conditions on the aggregate reports thru $\psi(\alpha)$ to identify the scaling factor α
- Our primary condition is independence between the total exposures $\sum_i w_i(d)$ and study day d.
- To satisfy this condition, we establish the following system $\psi(\alpha)$

$$\frac{n_k(1)}{n_k(d)} - \alpha_k(d) = 0, \forall k \tag{4}$$

where $n_k(d)$ is the number of observed reports on the dth study day and for the kth stratum (allowing for different attrition effects by subject or contextual factors).

Scaling Factor (2)

- A non-parametric approach to estimate α would simply solve for Eq. (4) using the observed $n_k(d)$.

- We can also consider a model to estimate n_k. A general form for the fatigue model,

$$g_0(n_k(d)) = g_1(d) + \epsilon \tag{5}$$

for some functions g_0 and g_1.

- Here, the fatigue model is only a function of d.
- Factors influencing the fatigue rate can be incorporated with stratification.

Model Selection

We propose using the following goodness-of-fit statistic to choose the model for estimating $\alpha(t)$

$$X^2 = \sum_{d=1}^{D} \sum_{k} \frac{(n_k(d) - \hat{n}_k(d))^2}{\hat{n}_k(d)}. \tag{6}$$

Smaller values of X^2 indicate a better fit to the observed fatigue pattern and a chi-squared test can be used to identify the most parsimonious model with adequate fit.

Over-Reporting

Over-reporting refers to reporting of alcohol advertisements during the collection period that would not have been observed had the student not been a participant in the study.

- We were concerned about possible over-reporting on the first days of the study
- One of the advantages of a parametric model for fatigue effects is that we can use the pattern in later study days to extrapolate reporting levels on the earliest study days and investigate the possibility of reports in excess of what would be expected

Assumptions

The proposed response model makes **two key assumptions**:

1. Ad reports are missing at random

2. There is a proportional relationship between the probability of compliance and the probability of event reporting, conditional on subject and contextual factors

Inference

- Parameters of interest (e.g. exposure rates, associations with exposure, etc.) are estimated with Horvitz-Thompson weighted estimators using weights $\hat{w}_i(t)$
- Confidence intervals and standard errors are obtained with bootstrap resampling methods

Predictors of Compliance

Given the nearly constant compliance rate for the TRAC study, we fit a linear model to the participant's average compliance to identify predictors of compliance.

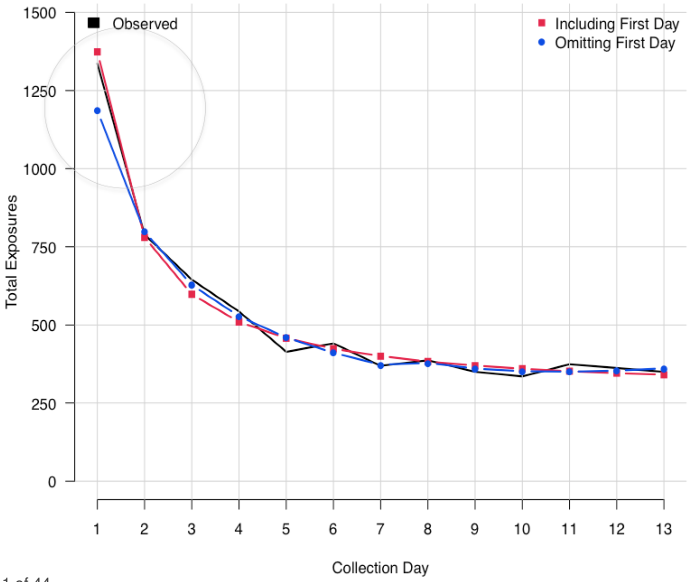
Factor	Estimate	P-value
Base Rate	72%	
Race		
White, non-Hispanic	(Ref)	
Hispanic	-5.8%	0.03
Black, non-Hispanic	-8.1%	0.01
Other	-0.8%	0.99
Grades in school	3.3%	0.01
Between-sibling random effect	17.1%	

Models for Scaling Factor

Model	X^2	P-value
Log-log (linear)	41.8	< 0.01
Log-log (quadratic)	18.5	0.05
Loess	31.6	< 0.01
Inverse	26.4	< 0.01
Log-log (linear), Stratified by weekend	44.8	< 0.01
Log-log (linear), Stratified by time of day	40.5	< 0.01
Log-log (linear), Stratified by week	13.6	0.19
Log-log (quadratic), Stratified by weekend	29.7	< 0.01
Log-log (quadratic), Stratified by time of day	18.8	< 0.01
Log-log (quadratic), Stratified by week	12.1	0.28
Log-log (linear) (-1)	36.4	< 0.01
Log-log (quadratic) (-1)	17.2	0.07
Inverse (-1)	24.7	< 0.01
Log-log (linear), Stratified by week (-1)	12.2	0.27
Log-log (quadratic), Stratified by week (-1)	11.3	0.18

(-1) Denotes models where the observed count on the first day was omitted

Evidence of Over-reporting



Description of Weights

Study Day	$\bar{\beta}$	α	w
1	1.51	0.89	1.34
2	1.51	1.49	2.25
3	1.52	1.89	2.86
4	1.51	2.25	3.40
5	1.52	2.58	3.91
6	1.53	2.89	4.43
7	1.52	3.18	4.84
8	1.53	3.14	4.79
9	1.51	3.29	4.97
10	1.52	3.37	5.11
11	1.53	3.38	5.16
12	1.53	3.35	5.11
13	1.52	3.28	4.99

Application: Daily Ad Exposure Rates

Characteristic	Observed (Unweighted)	Estimated (95% CI)
Overall	0.88	3.45 (3.42, 3.49)
Venue		
Outdoor	0.34	1.33 (1.28, 1.37)
Television	0.22	0.90 (0.86, 0.94)
Indoor	0.08	0.30 (0.28, 0.33)
Print	0.07	0.25 (0.22, 0.27)
Radio	0.05	0.20 (0.18, 0.22)
Online	0.04	0.13 (0.11, 0.15)
Product	0.03	0.12 (0.10, 0.14)
Item	0.03	0.11 (0.09, 0.12)
Other	0.02	0.06 (0.05, 0.08)

Summary

- EMA is an increasingly popular approach for the collection of repeated events.
- Because of the intensive monitoring used with EMA, non-reporting is common and efforts are needed at both the design and analysis stage to reduce non-reporting bias.
- We have presented a modeling approach for non-reporting that uses a two-stage model for missing event reports and inverse probability weighting to correct for missing reports in EMA applications.