Published in final edited form as:

Biometrics. 2016 March ; 72(1): 46–55. doi:10.1111/biom.12416.

Mixed Effects Models for Recurrent Events Data with Partially Observed Time-Varying Covariates: Ecological Momentary Assessment of Smoking

Stephen L. Rathbun and

Department of Epidemiology and Biostatistics, University of Georgia, Athens, Georgia, U.S.A.

Saul Shiffman

Department of Psychology, University of Pittsburgh, Pittsburgh, Pennsylvania, U.S.A.

Stephen L. Rathbun: rathbun@uga.edu; Saul Shiffman: shiffman@pinneyassociates.com

Summary

Cigarette smoking is a prototypical example of a recurrent event. The pattern of recurrent smoking events may depend on time-varying covariates including mood and environmental variables. Fixed effects and frailty models for recurrent events data assume that smokers have a common association with time-varying covariates. We develop a mixed effects version of a recurrent events model that may be used to describe variation among smokers in how they respond to those covariates, potentially leading to the development of individual-based smoking cessation therapies. Our method extends the modified EM algorithm of Steele (1996) for generalized mixed models to recurrent events data with partially observed time-varying covariates. It is offered as an alternative to the method of Rizopoulos, Verbeke and Lesaffre (2009) who extended Steele's (1996) algorithm to a joint-model for the recurrent events data and time-varying covariates. Our approach does not require a model for the time-varying covariates, but instead assumes that the time-varying covariates are sampled according to a Poisson point process with known intensity. Our methods are well suited to data collected using Ecological Momentary Assessment (EMA), a method of data collection widely used in the behavioral sciences to collect data on emotional state and recurrent events in the every-day environments of study subjects using electronic devices such as Personal Digital Assistants (PDA) or smart phones.

Keywords

Fully exponential Laplace approximation; Modified EM algorithm; Probability sample; Random covariate effects

0.1 Introduction

Cigarette smoking is a major cause of morbidity and mortality (U.S. Dept HHS 2014), making understanding of smoking behavior an urgent concern. While smoking is thought to

Supplementary Materials

Web Appendices referred to in Sections 2–4, simulated data and computer code are available with this paper at the *Biometrics* website on Wiley Online Library.

be driven by smokers' needs to regulate plasma nicotine levels (Schadel et al. 2000), the smoking of individual cigarettes may be cued or suppressed by situational factors that vary over time including mood (e.g., negative affect, restlessness), external environmental stimuli (e.g., others smoking, being alone) and smoking restrictions (Shiffman et al. 2002). Moreover, variation among smokers in how situational factors predict patterns of smoking prior to quit attempts, may be predictive of not only their chance of abstinence (Shiffman et al. 2007), but also the cues that may lead to relapse or resuming smoking. Mixed effects models for the impact of time-varying covariates on recurrent smoking events data may be used to describe this variation among subjects, potentially leading to the development of individually-tailored smoking cessation therapies (Strecher 1999).

Ecological Momentary Assessment (EMA; Stone & Shiffman, 1994; Shiffman et al. 2008) is a method in the behavioral sciences focused on collection of data about subjects' current psychological states in their everyday environments, frequently using electronic devices such as Personal Digital Assistants (PDAs) and smart phone APPs. EMA is particularly well suited for patient-reported outcomes related to subjective symptoms, quality of life, and emotional states (Hufford and Shiffman 2003). It has often been used to investigate recurrent addictive behavioral events such as the use of tobacco (Shiffman 2005), cocaine (Freedman et al. 2006), and alcohol (Collins et al. 1998). By assessing subjects' current states, EMA reduces recall bias inherent to retrospective assessments (Robinson and Clore 2002), and since the devices are carried by subjects throughout their normal daily activities, the ecological validity of the data is ensured (Shiffman and Stone 1998), meaning that the outcomes represent what happens in the everyday lives of the study subjects (Brewer 2000).

This paper considers mixed effects versions of point process models for an EMA of smoking behavior of 304 smokers who were attempting to quit (Shiffman et al. 2002). Subjects were asked to record each cigarette on a PDA programmed for data collection. The PDA prompted the smokers to answer questions regarding their current mood and environment at randomly selected cigarettes and also at randomly selected points in time from a known probability-based sampling design, resulting in a total of 15,262 event-contingent and 17,453 random assessments over the 13-day study interval.

For mixed effects versions of recurrent events models, the marginal distribution of the recurrent events is obtained by integrating the joint distribution of the events and the random effects with respect to the random effects, a requirement shared by generalized linear mixed models (GLMMs). Except for models with gamma frailties (e.g., Lawless 1987; Thall 1988), closed-form expressions for the marginal distribution of the recurrent events have not been obtained. While numerical integration might be used to approximate the marginal likelihood (e.g., Pinheiro and Bates 1995), numerical integration is not suitable for more than two or three random effects owing to the 'curse of dimensionality' (Bellman 1957). Laplace approximations to the likelihood (e.g., Breslow and Clayton 1993) and maximum hierarchical likelihood (Lee and Nelder 1996) are computationally efficient. The latter has been extended to frailty models for recurrent events data with time-varying covariates by Ha et al. (2001). Ha and Lee (2003) briefly consider hierarchical likelihood estimation of random covariate effects, but do not consider time-varying covariates. These various approximations yield inconsistent estimators when the sampling domain (period of time over

which each subject is sampled) is small, and the size the bias can be substantial if the variance components are not small (Breslow and Lin 1995, 1996). Kuk's (1999) Laplace importance sampling algorithm may be adapted to analysis of recurrent events data, but using simulations, Cole et al. (2003) find that it can also yield biased estimates. Random effects may be treated as missing data, so the EM algorithm could be applied to obtain consistent estimators for the mixed model parameters. McCulloch (1997) and Booth and Hobert introduce Monte Carlo versions of the EM algorithm that can yield consistent estimators, but these methods are computationally intensive and so may not be suitable for analysis of large EMA data sets.

We extend Steele's (1996) modified EMA algorithm to the analysis of recurrent events data with partially observed time-varying covariates. This algorithm is computationally efficient, and uses the second-order Laplace approximation of Tierney et al. (1989) in the M-step, leading to a better approximation to the marginal likelihood when compared to approaches based on a first-order approximation. Moreover, computer simulations suggest that resulting estimators have smaller bias than competing techniques (Steele 1996; Cole et al. 2003).

Likelihood-based inference for the impact of time-varying covariates on recurrent events data requires the evaluation of the integrated intensity and its derivatives with respect to model parameters, quantities that are unknown unless covariates are observed functions of time. To model to impact of partially observed time-varying covariates on the lifetimes to events, Wulfson and Tsiatis (1997) construct a joint model for the time-varying covariates and the event lifetimes in the context where only a single lifetime is observed for each subject. Liu and Huang (2007) and Zhang et al. (2008) extend this joint modeling approach to recurrent events data. Rizopoulos et al. (2009) extended Steele's (1996) approach to joint models for survival and longitudinal data. Although a wide variety of models are available for time-varying covariates including polynomial, spline and stochastic process models (Tsiatis and Davidian 2004), estimators of covariate effects may be biased if the model for the covariates as a function of time is misspecified (Rathbun et al. 2013). The sampling path of mood as a potential predictor of smoking patterns (Kassel et al. 2003; Shiffman et al. 2002) is not a continuous function of time as it may change abruptly in response to multiple stimuli. Since such stimuli may occur at frequencies greater than the frequency at which mood is sampled, no model is likely to adequately capture this variation in mood, so the joint modeling approach may yield biased estimates of point process model parameters.

Rathbun et al. (2007) propose an approach for modeling the impact of partially-observed time-varying covariates on Poisson process intensity that does not involve a model for the time-varying covariates. A design-unbiased estimator (Cassel et al. 1977) for the integrated intensity and its derivatives is constructed from the probability sample of the covariates and substituted into the likelihood equations to obtain an estimator for the point process parameters. Waagepetersen (2008) improved upon this approach by using covariate information from the events as well as the probability sample to construct a design-unbiased estimator for the integrated intensity. Estimating equations of both Rathbun et al. (2007) and Waagepetersen (2008) belong to a general class of weighted estimating equations. Rathbun (2013) demonstrated that provided that the covariates are sampled from an inhomogeneous Poisson process with known intensity $\pi(t)$ Waagepetersen's estimating equations are

optimal (sensu Ferreira 1982) within this class of weighted estimating equations. Moreover, Waagepetersen's estimating equations are equivalent to the logistic regression score equations for the binary random variable taken to be one if the observation is an event and zero if it is from the probability sample with offset $-\log \pi(t)$. Although Rathbun et al. (2007) and Waagepetersen (2008) are only concerned with Poisson point processes, the methods of Rathbun et al. (2013) can be used to extend their results to point processes defined through conditional intensity functions depending on the past history of recurrent events.

In this paper, we will employ the probability generating functional to demonstrate that the logistic regression likelihood well approximates the point process likelihood suggesting that Steele's (1996) fully exponential Laplace approximation, intended for GLMMs may also be used to fit mixed effects versions of point process models to recurrent events data. Under the estimating equations of Rathbun et al. (2007), and the approach proposed in the current paper covariates are treated as deterministic functions of time, so statistical inference is carried out conditional on the realization of the sample paths of the time-varying covariates.

In Section 2, we describe the mixed effects model for recurrent events data with time-varying covariates, and demonstrate that provided that the recurrent events and probability sample of covariates are realized from independent point processes, the full-data likelihood may be approximated by an expression equivalent to the full-data likelihood in a mixed effects version of logistic regression. Section 3 extends the modified EM algorithm of Steele (1996) to the current application, including fully exponential Laplace approximations for the expectations in the E-step of the algorithm. Issues regarding implementation of the proposed algorithm, its convergence and large-sample properties are discussed. The results of a simulation study are described in Section 4, and the application of the proposed methods to an EMA of smoking are presented in Section 5.

1. Mixed Effects Model for Recurrent Events

For a mixed effects model for recurrent smoking events, assume that conditional on the realization of a vector of random effects β_i , the event times for each subject i are realized from a point process with conditional intensity

$$\lambda_i(t; \alpha, \beta_i) = \exp\{\alpha^T \mathbf{z}_i(t) + \beta_i^T \tilde{\mathbf{z}}_i(t)\}; t \in A_i, i=1, \dots, n, \quad (1)$$

where the random effects $\tilde{\mathbf{z}}_i(t)$ are a subset of the fixed effects $\mathbf{z}_i(t)$, $A_i \subset \mathbb{R}$ is the sampling domain for subject i and where one or more elements of $\mathbf{z}_i(t)$ may depend on the history of past events. One limitation of our approach is that only fully parametric models may be fit to the data. However, proportional hazards models may be considered, for example, by fitting spline models to the log baseline hazard, in which case some of the elements of $\mathbf{z}_i(t)$ are spline basis functions. The random effects β_i are assumed to be independently sampled from a zero-mean normal distribution with variance-covariance matrix Σ . The parameters α describe the impact the covariates on smoking risk in the population, while the random effects β_i describe how the impact of those covariates varies among smokers. Positive random effects β_{ik} indicate that the corresponding covariate $\tilde{z}_{ik}(t)$ puts smokers at increased

risk of smoking a cigarette relative to the population, while negative random effects indicate decreased risk. The sets A_j may be a single interval of time, but in the current application, A_j is the union of a finite number of disjoint intervals corresponding to the set of times during which the PDA is actively collecting data. The PDA was suspended during sleep and naps and when the subject turns it off for meetings, movies, etc. For each subject, let $X_i = \{t_{ij} : j = 1, \dots, N_i\}$ denote the set of recurrent event times.

Conditional on the random effects, the point process X_i has Janossy density

$$L_i(X_i; \alpha, \beta_i) = \exp \left\{ \alpha^T \sum_{t \in X_i} \mathbf{z}_i(t) + \beta_i^T \sum_{t \in X_i} \tilde{\mathbf{z}}_i(t) - \Lambda_i(A_i; \alpha, \beta_i) \right\} \quad (2)$$

(Daley and Vere-Jones 1988), where

$$\Lambda_i(A_i; \alpha, \beta_i) = \int_{A_i} \exp \{ \alpha^T \mathbf{z}_i(t) + \beta_i^T \tilde{\mathbf{z}}_i(t) \} dt$$

is the integrated intensity. Let $\phi_i(\beta_i; \Sigma)$ denote the probability density function for a zero-mean normally distributed random vector with variance-covariance matrix Σ . Then the contribution of subject i to the marginal distribution of the data is

$$L_i(X_i; \alpha, \Sigma) = \int_{\mathbb{R}^q} L_i(X_i; \alpha, \beta_i) \phi_i(\beta_i; \Sigma) d\beta_i.$$

This marginal distribution is a function of the Janossy density (2), a function of $\exp \{-\Lambda_i(A_i; \alpha, \beta_i)\}$, the evaluation of which requires that the covariates be known functions of time. In most practical applications, however, covariates may only be sampled over time. For EMA, covariates are sampled at times t belonging to a random set $D_i \subset A_i$ selected according to a known probability-based sampling design. For each (α, β_i) , $\Lambda_i(A_i; \alpha, \beta_i)$ may be regarded as a population total for $\lambda_i(t; \alpha, \beta_i)$, where the population is the collection of points $t \in A_i$. Suppose that for each subject i , covariates are assessed at times t sampled from a point process D_i with known (conditional) intensity $\pi_i(t)$. For fixed effects models, Waagepetersen (2008) replaces $\Lambda_i(A_i; \alpha)$ with a design-unbiased estimator constructed using covariate data from the superposition of the recurrent event process X_i and the probability sample D_i . Since the Poisson score equations are linear in $\Lambda_i(A_i; \alpha)$, a consistent estimator for the fixed effects is obtained. Since it is nonlinear in $\Lambda_i(A_i; \alpha, \beta_i)$, simply substituting a design-unbiased estimator for $\Lambda_i(A_i; \alpha, \beta_i)$ yields a biased estimator for $\exp \{-\Lambda_i(A_i; \alpha, \beta_i)\}$. However, insight into the design-unbiased estimation of $\exp \{-\Lambda_i(A_i; \alpha, \beta_i)\}$ can be obtained from the probability generating functional.

For functions $\xi(t)$ bounded between zero and one, define the probability-generating functional of a point process N^* on A with conditional intensity $\zeta(t)$

$$G_N(\xi) = E \left[\exp \left\{ \sum_{t \in X^*} \log \xi(t) \right\} \right]$$

(Vere-Jones 1968; Westcott 1972). If X^* is a Poisson point process, then it is well known that the probability generating functional is

$$G_N(\xi) = \exp \left[- \int_A \{1 - \xi(t)\} \zeta(t) dt \right] \quad (3)$$

(e.g., Daley and Vere-Jones 1988, p. 225). In Web Appendix A, we demonstrate that the above result also holds when N^* has conditional intensity $\zeta(t)$, depending on the history of past events. If we take

$$\xi_i(t; \alpha, \beta_i) = \frac{\pi_i(t)}{\pi_i(t) + \lambda_i(t; \alpha, \beta_i)}, \quad (4)$$

and $X_i^* = X_i \cup D_i$ to be the superposition of the event process with intensity $\lambda_i(t; \alpha, \beta_i)$ and covariate sampling process with intensity $\pi_i(t)$, then conditional on the realization of the random effects,

$$E \left[\exp \left\{ \sum_{t \in X_i \cup D_i} \log \frac{\pi_i(t)}{\pi_i(t) + \lambda_i(t; \alpha, \beta_i)} \right\} \middle| \beta_i \right] = \exp \{ -\Lambda_i(A_i; \alpha, \beta_i) \},$$

as desired. This suggests that $\Lambda_i(A_i; \alpha, \beta_i)$ in the Janossy density may be replaced by $\log \xi_i(t; \alpha, \beta_i)$ to yield the approximate Janossy density

$$\hat{L}_i(X_i; \alpha, \beta_i) = \exp \left[\alpha^T \sum_{t \in X_i} \mathbf{z}_i(t) + \beta_i^T \sum_{t \in X_i} \tilde{\mathbf{z}}_i(t) - \sum_{t \in N_i \cup D_i} \log \{1 + \lambda_i(t; \alpha, \beta_i) / \pi_i(t)\} \right].$$

Note that this is the logistic regression likelihood for the binary random variable $Y_i(t)$, where $Y_i(t) = 1$ for $t \in X_i$ and $Y_i(t) = 0$ for $t \in D_i$ and offset $-\log \pi_i(t)$.

2. Modified EM Algorithm

The above arguments suggest that the fully exponential Laplace approximation Steele (1996) proposed for GLMMs may readily be extended to analysis of recurrent events data, where covariates are sampled according a point process with known intensity as well as at the event times. Parameters are estimated according to a modified version of the EM algorithm, where conditional expectations are approximated using a second-order Laplace approximation.

Iterate u of the E-step of the EM algorithm involves the computation of the conditional expectation of the complete-data log likelihood, which in the current context is

$$\hat{Q}(\boldsymbol{\theta}|\boldsymbol{\theta}^{(u)}) = \sum_{i=1}^n \left[\left\{ \boldsymbol{\alpha}^T \sum_{t \in X_i} \mathbf{z}_i(t) + E^{(u)}(\boldsymbol{\beta}_i^T) \sum_{t \in X_i} \tilde{\mathbf{z}}_i(t) - E^{(u)} \left(\sum_{t \in X_i \cup D_i} \log[1 + \lambda_i(t)/\pi_i(t)] \right) \right\} - \frac{1}{2} E^{(u)}(\boldsymbol{\beta}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta}_i) \right] - \frac{n}{2} \log|\boldsymbol{\Sigma}|,$$

where the dependence of $\lambda_i(t)$ on $(\boldsymbol{\alpha}, \boldsymbol{\beta}_i)$ is dropped for notational convenience, and the expectations $E^{(u)}(\cdot)$ are conditional on the event process X_i , the covariate sampling process D_i , and the current estimates of the parameters $\boldsymbol{\theta}^{(u)} = (\boldsymbol{\alpha}^{(u)}, \boldsymbol{\gamma}^{(u)}, \boldsymbol{\Sigma}^{(u)})$. In the M-step, the parameters $\boldsymbol{\alpha}$ and $\boldsymbol{\Sigma}$ are updated by finding $\boldsymbol{\theta}^{(p+1)}$ that maximizes $Q(\boldsymbol{\theta}|\boldsymbol{\theta}^{(p)})$. For the fixed effects $\boldsymbol{\alpha}$, this entails finding $\boldsymbol{\alpha}^{(u+1)}$ that solves the estimating equations

$$\mathbf{S}(\boldsymbol{\alpha}) = \sum_{i=1}^n \left[\sum_{t \in X_i} \mathbf{z}_i(t) - E^{(u)} \left\{ \sum_{t \in X_i \cup D_i} \{1 - \xi_i(t)\} \mathbf{z}_i(t) \right\} \right] = 0, \quad (5)$$

where $\xi_i(t) = \pi_i(t) [\pi_i(t) + \lambda_i(t)]^{-1}$. To update the estimates of $\boldsymbol{\Sigma}$, note that $E^{(u)}(\boldsymbol{\beta}_i^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta}_i) = \text{trace}\{\boldsymbol{\Sigma}^{-1} E^{(u)}(\boldsymbol{\beta}_i \boldsymbol{\beta}_i^T)\}$, so by Lemma (3.2.2) of Anderson (1984), $\boldsymbol{\Sigma}$ may be updated using

$$\boldsymbol{\Sigma}^{(u+1)} = \frac{1}{n} \sum_{i=1}^n E^{(u)}(\boldsymbol{\beta}_i \boldsymbol{\beta}_i^T). \quad (6)$$

2.1 Fully-Exponential Laplace Approximation

The E-Step of the EM algorithm requires the computation of conditional expectations of functions of the general form $g(\boldsymbol{\beta}_i)$

$$E\{g(\boldsymbol{\beta}_i)\} = \frac{\int g(\boldsymbol{\beta}_i) \exp\{|A_i| h_i(\boldsymbol{\beta}_i)\} d\boldsymbol{\beta}_i}{\int \exp\{|A_i| h_i(\boldsymbol{\beta}_i)\} d\boldsymbol{\beta}_i}, \quad (7)$$

where $h_i(\boldsymbol{\beta}_i) = |A_i|^{-1} \{\log L(\hat{X}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}_i) + \log \phi_i(\boldsymbol{\beta}_i; \boldsymbol{\Sigma})\}$ and $|A_i|$ is the total length of time over which subject i is observed. For GLMMs, the complete data log likelihood is normalized by the number of repeated observations n_i of each subject i , and a second-order Laplace approximation yields an error of order $O(n_i^{-1})$. In our context, however, we observe a random number N_i of recurrent events and a random number M_i of covariate samples for each subject, both of which are of order $O_p(|A_i|)$. So, normalizing by $|A_i|^{-1}$ is appropriate, and a second-order Laplace approximation yields an error of order $O_p(|A_i|^{-1})$. Assuming that $g(\boldsymbol{\beta}_i)$ is strictly positive, applying the Laplace approximation to both the numerator and denominator of (7) leads to the cancellation of the $O_p(|A_i|^{-1})$ terms and so, yields an approximation of order $O_p(|A_i|^{-2})$ for the conditional expectations.

The conditional expectations in the expressions (5), and (6) are of terms that are not necessarily positive. Following the suggestions of Steele (1996) and Rizopoulos et al. (2009), we apply a multivariate version of the fully exponential Laplace approximation of

Tierney et al. (1989) for the conditional cumulant generating function $\log E[\exp \{\mathbf{s}^T \mathbf{g}(\boldsymbol{\beta}_i)\} | X_i, \boldsymbol{\alpha}, \boldsymbol{\Sigma}]$. Then the conditional expectation can be obtained by taking the derivative of the cumulant generating function with respect to $\mathbf{s}$ evaluated at $\mathbf{s} = \mathbf{0}$.

For each subject i , let $\boldsymbol{\beta}_i^*(\mathbf{s})$ denote the mode of $\log L(\hat{X}_i; \boldsymbol{\alpha}, \boldsymbol{\beta}_i) + \log \varphi_i(\boldsymbol{\beta}_i; \boldsymbol{\Sigma}) + \mathbf{s}^T \mathbf{g}(\boldsymbol{\beta}_i)$ and define

$$\mathbf{V}_i(\mathbf{s}, \boldsymbol{\beta}_i) = -\frac{\partial^2}{\partial \boldsymbol{\beta}_i \partial \boldsymbol{\beta}_i^T} \left\{ \log \hat{L}(X_i; \boldsymbol{\alpha}, \boldsymbol{\beta}_i) + \log \varphi_i(\boldsymbol{\beta}_i; \boldsymbol{\Sigma}) + \mathbf{s}^T \mathbf{g}(\boldsymbol{\beta}_i) \right\} = \sum_{t \in X_i \cup D_i} w_i(t; \boldsymbol{\beta}_i) \tilde{\mathbf{z}}_i(t) \tilde{\mathbf{z}}_i^T(t) + \boldsymbol{\Sigma}^{-1} - \mathbf{s}^T \frac{\partial^2}{\partial \boldsymbol{\beta}_i \partial \boldsymbol{\beta}_i^T} g(\boldsymbol{\beta}_i), \quad (8)$$

where $w_i(t; \boldsymbol{\beta}_i) = \xi_i(t; \boldsymbol{\beta}_i) \{1 - \xi_i(t; \boldsymbol{\beta}_i)\}$. Then the second-order Laplace approximation for element m of $E^{(u)} \{\mathbf{g}(\boldsymbol{\beta}_i)\}$ is

$$\hat{E}^{(u)} \{g_m(\boldsymbol{\beta}_i) | X_i\} = g_m\{\boldsymbol{\beta}_i^*(\mathbf{0})\} + A[g_m\{\boldsymbol{\beta}_i^*(\mathbf{0})\}], \quad (9)$$

where $\hat{E}^{(u)} \{g_m(\boldsymbol{\beta}_i) | X_i\} = E^{(u)} \{g_m(\boldsymbol{\beta}_i) | X_i\} + O_p(|A_i|^{-2})$ and

$$A[g_m\{\boldsymbol{\beta}_i^*(\mathbf{0})\}] = -\frac{1}{2} \text{trace} \left\{ \mathbf{V}_i(\mathbf{0}, \boldsymbol{\beta}_i^*(\mathbf{0}))^{-1} \frac{\partial}{\partial s_m} \mathbf{V}_i(\mathbf{s}, \boldsymbol{\beta}_i^*(\mathbf{s})) \Big|_{(\mathbf{s}, \boldsymbol{\beta}_i) = (\mathbf{0}, \boldsymbol{\beta}_i^*(\mathbf{0}))} \right\}.$$

is the adjustment factor for element m of $\mathbf{g}(\boldsymbol{\beta}_i)$ evaluated at $\boldsymbol{\beta}_i = \boldsymbol{\beta}_i^*(\mathbf{0})$. To obtain a more explicit expression for the adjustment $A[g_m\{\boldsymbol{\beta}_i^*(\mathbf{0})\}]$, define the first and second order derivatives $D_m g_m^* = (\partial / \partial \boldsymbol{\beta}_i) g_m(\boldsymbol{\beta}_i) |_{\boldsymbol{\beta}_i^*(\mathbf{0})}$ and $D_m^2 g_m^* = (\partial^2 / \partial \boldsymbol{\beta}_i \partial \boldsymbol{\beta}_i^T) g_m(\boldsymbol{\beta}_i) |_{\boldsymbol{\beta}_i^*(\mathbf{0})}$ of $g_m(\boldsymbol{\beta}_i)$ with respect to the random effects. Then compute

$$D_m w_i = \frac{\partial}{\partial s_m} w_i(t; \boldsymbol{\beta}_i^*(\mathbf{s})) \Big|_{(\mathbf{s}, \boldsymbol{\beta}_i) = (\mathbf{0}, \boldsymbol{\beta}_i^*(\mathbf{0}))} = -\zeta_i(t; \boldsymbol{\beta}_i^*(\mathbf{s})) \tilde{\mathbf{z}}_i^T(t) \mathbf{V}_i(\mathbf{0}, \boldsymbol{\beta}_i^*(\mathbf{0}))^{-1} D_m g_m^*,$$

for $t \in X_{ij} \cup D_{ij}$ where $\zeta_i(t; \boldsymbol{\beta}_i) = \xi_i(t; \boldsymbol{\beta}_i) [1 - \xi_i(t; \boldsymbol{\beta}_i)] [1 - 2\xi_i(t; \boldsymbol{\beta}_i)]$. Then the adjustment factor for element m of $E^{(u)} \{\mathbf{g}(\boldsymbol{\beta}_i)\}$ is

$$A[g_m\{\boldsymbol{\beta}_i^*(\mathbf{0})\}] = -\frac{1}{2} \text{trace} \left[\mathbf{V}_i^{*-1} \left\{ \sum_{t \in X_i \cup D_i} \zeta_i^*(t) \tilde{\mathbf{z}}_i(t) \tilde{\mathbf{z}}_i^T(t) (\tilde{\mathbf{z}}_i^T(t) \mathbf{V}_i^{*-1} D_m g_m^* - D_m^2 g_m^*) \right\} \right], \quad (10)$$

where $\zeta_i^*(t) = \zeta_i(t; \boldsymbol{\beta}_i^*(\mathbf{0}))$ and $\mathbf{V}_i^* = \mathbf{V}_i(\mathbf{0}, \boldsymbol{\beta}_i^*(\mathbf{0}))$.

2.2 Implementation of the Modified EM Algorithm

Implementation of iterate u of the E-Step of Steele's (1996) modified EM algorithm requires first the solution to the complete data score equations for the random effects,

$$\boldsymbol{\psi}_i(\boldsymbol{\beta}_i) = \sum_{t \in X_i} \tilde{\mathbf{z}}_i(t) - \sum_{t \in X_i \cup D_i} \{1 - \xi_i(t; \boldsymbol{\beta}_i)\} \tilde{\mathbf{z}}_i(t) - \boldsymbol{\Sigma}^{-1} \boldsymbol{\beta}_i = \mathbf{0}, \quad (11)$$

followed by the above second-order Laplace approximation for the conditional expectations in expressions (5) and (6). Taking β_i to be a fixed parameter, the solution $\beta_i^{(u)}$ to (11) may be regarded as a penalized likelihood estimator, shrinking estimators of the random effects towards their zero mean. While a closed-form expression for the solutions to (11) does not generally exist, an iterative solution may be obtained using the Newton-Raphson algorithm. Using $\beta_i^{(u)}$ as a starting value, estimates of $\beta_i^{(v)}$ may be updated by taking

$$\beta_i^{(v+1)} = \beta_i^{(v)} + \mathbf{V}_i \left(\mathbf{0}, \beta_i^{(v)} \right)^{-1} \psi_i \left(\beta_i^{(v)} \right),$$

where $\mathbf{V}_i \left(\mathbf{0}, \beta_i^{(v)} \right)$ is defined as in expression (8). Following the first few iterations of the EM algorithm, convergence is achieved within a few Newton-Raphson steps.

The E-step of the algorithm is completed by replacing the conditional expectations $E^{(u)}(\cdot)$ in expressions (5) and (6) with their respective second-order Laplace approximations $\hat{E}^{(u)}(\cdot)$ (9). Here, the vector $\mathbf{g}(\cdot)$ in (7) is taken to be comprised of the elements of

$\sum_{t \in X_i \cup D_i} \left(1 - \xi_i \left(t; \beta_i^{(u)} \right) \right) \mathbf{z}_i(t)$ in expression (5), and $\beta_i \beta_i^T$ in expression (6).

Computation of the second-order Laplace approximation $\hat{E}^{(u)}(\cdot)$ requires the computation of the adjustment factor (10) which in turn requires the computation of the first- and second-order derivatives of $\mathbf{g}(\cdot)$ as detailed in Web Appendix B. The former results in the construction of approximate score equations

$$\hat{S}(\alpha_j) = \sum_{i=1}^n \left[\sum_{t \in X_i} z_{ij}(t) - \sum_{t \in X_i \cup D_i} \left\{ 1 - \xi_i \left(t; \beta_i^{(u)} \right) \right\} z_{ij}(t) - A \left\{ \sum_{t \in X_i \cup D_i} \left\{ 1 - \xi_i \left(t; \beta_i^{(u)} \right) \right\} z_{ij}(t) \right\} \right] = 0; \quad (12)$$

for α_j , $j = 1, \dots, p$. Then the fixed effects may be updated using a single Newton-Raphson step $\alpha^{(u+1)} = \alpha^{(u)} + \{ \mathbf{H}(\beta^{(u)}) \}^{-1} \hat{\mathbf{S}}(\alpha^{(u)})$ where

$$\mathbf{H}(\alpha; \beta) = \sum_{i=1}^n \sum_{t \in X_i \cup D_i} w_i \left(t; \beta_i^{(u)} \right) \mathbf{z}_i(t) \mathbf{z}_i^T(t)$$

is the complete-data Hessian, $\alpha = (\alpha_1, \dots, \alpha_p)^T$, and $\hat{\mathbf{S}}(\alpha) = (\hat{S}(\alpha_1), \dots, \hat{S}(\alpha_p))^T$. The variance components are updated using

$$\Sigma^{(u+1)} = \frac{1}{n} \sum_{i=1}^n \left[\beta_i^{(u)} \left(\beta_i^{(u)} \right)^T + A \left\{ \beta_i^{(u)} \left(\beta_i^{(u)} \right)^T \right\} \right]. \quad (13)$$

Methods for predicting the random effects are similar and are detailed in Web Appendix C.

Convergence of the modified EM algorithm requires that the surrogate function $Q(\hat{\theta}|\theta^{(u)})$ or its approximation be an increasing function of the iterate u (Lang et al. 2000). By arguments similar to those presented by Rizopoulos et al. (2009), the continuity of $Q(\hat{\theta}|\theta^{(u)})$ and its approximations with respect to θ and $\theta^{(u)}$ is sufficient to ensure that $Q(\hat{\theta}|\theta^{(u)})$ is an

increasing function of u . However, the proposed algorithm can fail to converge given its use of Newton-Raphson steps unless good starting values for model parameters are used. To obtain starting values for α , we first fit a frailty model using the methods described in

Neustifter et al. (2012). For the variance components, we initialize $\Sigma = \text{diag} \left(\sqrt{|\alpha_{1k}|} \right)$.

As detailed in the Web Appendix D, the proposed estimator $\hat{\theta} = \theta_0 + O_p[\max\{n^{-1/2}, \min |A_{ji}|^{-2}\}]$ where θ_0 is the true value of the model parameter. The SEM algorithm of Meng and Rubin (1991) was used to estimate the asymptotic variance-covariance matrix of the parameter estimates. As suggested by Meng and Rubin, the symmetry of the resulting variance-covariance matrix is used as an informal diagnostic for the precision with which the variance-covariance matrix is estimating.

3. Simulated Data

As detailed in Web Appendix E, simulations were carried out to assess the small-sample properties of the proposed approach under a Poisson point process, and point processes with negative and positive dependence among recurrent events. Each model consisted of a single fixed effect and two random effects as well as a random intercept term to describe variation among subjects in baseline smoking rates. Simulations were carried out over a 13-day period to match the interval over which data were collected in the smoking study. Samples of $n = 100$ and $n = 300$ subjects were considered and covariates were sampled with intensities of $\pi = 2.0$ and $\pi = 4.0$ events per day. The proposed approach showed good performance even at the smaller sample sizes and covariate sampling intensities. Variances of random effects tended to be modestly over-estimated with biases up to 7% and covariances between the two random effects were under-estimated. Nevertheless, true values of each element of the variance covariance matrix lied well within 95% intervals defined the the 2.5th and 97.5th percentiles. Estimates of the mean of the random effects and the fixed effect showed little bias, with the random effect having the higher variance showing the most bias, and fixed effect showing the least bias. Moreover, coverage rates of nominal 95% confidence intervals generally exceeded 95%, especially under the Poisson process, a consequence of conservative estimates of standard errors obtained under the SEM algorithm. Less conservative estimates of standard errors could be achieved using the complete-data information matrix, but this would result in coverage rates falling short of nominal 95% levels. Standard errors were largest under positive dependence among recurrent events, and smallest under negative dependence; the Poisson point process yielded intermediate standard errors.

4. Ecological Momentary Assessment of Smoking

The following illustrates mixed effects modeling of recurrent events using data from an EMA of smoking. A total of 304 smokers expressing a desire to quit were trained to use PDAs programed for data collection. Smokers were instructed to continue smoking according to their usual habits for 16 days prior to a designated quit date, and to record each cigarette smoked on the PDA. The first three days of monitoring were not considered to yield reliable data since subjects needed time to acclimate to the use of the PDA, and so

these data were dropped from the analysis. Covariate information on mood and environment was obtained from electronically administered assessments at the times of randomly selected cigarettes and at times selected from a random-sampling design. Environmental variables included smoking restrictions (allowed, discouraged, and prohibited), presence/absence of other smokers, and whether or not the smoker was alone. Participants rated mood adjectives derived from the circumplex model of affect (Russell 1980) with items scored on a 4-point Likert scale. Factor analysis yielded three orthogonal factors (Shiffman et al. 1996), negative affect, arousal, and attention disturbance. Loadings for the item 'restlessness' were small for all three factors, and so this item was retained as a separate covariate in subsequent analyses.

On each day, interarrival times of covariate samples were generated from a uniform distribution over a 2-hour period, initiated at the wake-up times of the subjects. Unfortunately, covariate sampling intensities $\pi_i(t)$ were not recorded and owing to missing data could not be reconstructed for all covariate samples. Therefore, we set the sampling intensities to the empirical rates $\pi_i(t) = m_{ij}/|A_{ij}|$ for $t \in A_{ij}$ where m_{ij} is the number of times that covariates were sampled on day j for smoker i and A_{ij} is the set of times the diary was active for smoker i on day j . To reduce the burden on study subjects, not all cigarettes were assessed for the covariates. To target five cigarette assessments per day, smoking events for subject i on day j were independently selected to be assessed with known probability

$p_{ij} = \min\{5N_{i,j-1}^{-1}, 1\}$, where $N_{i,j-1}$ is the number of smoking events for subject i on day $j-1$. Conditional on $N_{i,j-1}$, and the random effect β_i , the thinned point process of assessed events is realized from a process with intensity $p_{ij} \exp\{\alpha^T \mathbf{z}_i(t) + \beta_i^T \tilde{\mathbf{z}}_i(t)\}$, or equivalently including an offset of $\log(p_{ij}/\pi_i(t))$ to the logistic regression model.

A mixed-effects model was fit to the data, predicting smoking intensity as a function of the two mood variables negative affect and restlessness, and three environmental variables alone, other smokers and smoking restrictions. All of the above variables including the intercept were treated as random effects, except for smoking restrictions which were treated as fixed effects.

Table 1 compares the results of fitting the mixed effects model to the fixed effects model fit using logistic regression as justified by Rathbun (2013), and the frailty model fit using the methods of Neustifter et al. (2012). Under the frailty model, the intensity function for subject i is $u_i \exp\{\alpha^T \mathbf{z}_i(t)\}$, where the frailty u_i is sampled from any distribution with mean one and variance σ^2 . Neustifter's estimator for the regression coefficients under the frailty model are equal to those of the fixed-effects model, but standard errors are substantially increased under the frailty model. The variance of the frailty is estimated to be 0.647. Except for smoking prohibitions which were treated as fixed effects, the estimates of the coefficients under mixed effects model are estimates of the mean of the corresponding random effects. Comparing these to the corresponding fixed effects estimates, no consistent pattern emerges, some are higher and some lower. This was also true for the standard errors.

Table 2 shows the estimated risk ratios for the mean impact of negative affect and restlessness under the mixed-effects model, computing by taking exponential function of the parameter estimates. After adjusting for the effects of the remaining predictors, smoking rate

was estimated to increase by 30.5% for every unit in restlessness, and 10.4% for each unit increase in negative affect. The former is consistent with our previous findings (Shiffman et al. 2002; Rathbun et al. 2007), while that latter had not been observed in our previous analysis that had not controlled for the environmental variables. Restlessness may reflect the undirected activation of motivational systems that drive towards smoking (Shiffman and Waters 2004). Previous analyses based on GEE had shown no effect of Negative Affect (Shiffman et al. 2002), which was contrary to commonly-accepted theoretical models (Kassel et al. 2003), so this is a new finding that supports theory and is consistent with smokers' overall self-reports (Kassel et al. 2003). Negative affect may increase both the motivation to relieve distress and undermine smokers' abilities to cope with temptation because experiencing emotional distress makes it difficult to call up coping resources. Importantly, the effect of negative affect on smoking may differ across smokers; prior analyses of EMA data have suggested differences by gender (Rathbun & Shiffman, 2011), and a relationship between individual differences in negative affect smoking and relapse risk (Shiffman et al. 2007).

The variables reflecting social constraints on smoking showed particularly strong effects on smoking rates. After controlling for the remaining variables, smokers smoked an estimated 0.371 cigarettes per hour (cph) when smoking was prohibited, 0.944 cph when smoking was discouraged, and 1.424 cph when smoking was allowed. Smoking rates were also higher when other smokers were present and when the smoker was alone. Others' smoking may provide visual, olfactory or social cues encouraging a smoker to light up (Shiffman et al. 2002). Laboratory studies have shown that the sight or smell of cigarettes (Sayette and Hufford 1994) and the sight of someone smoking can elicit craving for cigarettes (Warren and McDonough 1999). Conversely, the presence of others who are not smoking may suppress smoking, out of concern for others' preferences.

Estimated standard deviations of random effects were large relative to their estimated means suggesting that there is considerable variation among smokers' responses to time-varying mood and environment variables (Table 1). Figure 1 presents the distributions of the estimated random effects, centered on their respective means. To varying degrees, all smokers smoked more frequently when they were restless. While a majority of smokers tend to smoke more frequently when they were experiencing negative emotions, 38.8% of smokers were estimated to smoke more frequently when they were not experiencing negative emotions. Likewise, while being alone and being in the presence of other smokers appears to be linked to higher smoking rates among a majority of smokers, small minorities of smokers showed increasing smoking rates when with others (7.2%) and when not being in the presence of other smokers (3.3%). However, these small percentages may be associated with smokers for whom these variables have no impact on smoking rates; negative estimates of coefficients may simply be due to error.

Table 3 depicts the estimated correlations among the random effects. Although most of the correlations among the random effects were negligible, a modest positive correlation was observed between negative affect and restlessness suggesting that those who smoke when they are restless also tend to smoke when they are experiencing negative emotions. These individuals may experience restlessness as negative, or experience negative emotions and

restlessness at the same time. Correlations were also observed between the intercept and the two environmental covariates indicating the heavy smokers tend to smoke more frequently when they are alone and when they are not with other smokers.

A referee suggested that "evidence of departure from a fixed effect model could derive from non-linear associations of the covariate with smoking intensity, rather than because of variability in regression parameters among study subjects." This is not relevant to the binary variables alone and other smokers, but could apply to the continuous variables negative affect and restlessness if there is significant variation among smokers in the distribution of those variables. Using ANOVA, we did observe significant variation among smokers in mean levels of both negative affect and restlessness with $R^2 = 0.33$ for negative affect and $R^2 = 0.35$ for restlessness, so we cannot eliminate the possibility of non-linear effects. However, we are not aware of any statistical method for assessing the relative merits mixed-effects vs. nonlinear models, but find the mixed effects model to be more useful for addressing the potential for individual-based therapies for smoking cessation.

5. Discussion

Our approach is offered as an alternative to the joint modeling approach of Rizopoulos et al. (2009) whose method for fitting mixed effects models to recurrent events data can lead to biased estimates of model parameters when the covariate model is misspecified (Rathbun et al. 2013). Since we do not require a covariate model, our approach is particularly well suited to settings where the covariates are not a continuous function of time, but are subject to frequent extrinsic perturbations. Our approach is based on probability-based samples of the time-varying covariates and assumes that the sample paths of time-varying covariates are deterministic functions of time, so inferences are carried out conditional on the realization of the time-varying covariates. Although we require a fully parametric intensity function, the log baseline intensity can be modeled by using spline or other suitably flexible basis functions. Our approach also requires that the time-varying covariates be sampled according to a point process with known (conditional) intensity. Implementation with convenience samples can lead to biased estimates of model parameters when a representative sample is not obtained. The methods proposed here are particularly well suited to data collected using EMA, where the use of probability-based sampling designs for sampling time-varying covariates is routine (Stone et al. 2007; Shiffman 2007).

The referee pointed out that another limitation of our approach is that the intensity model (1) involves only current values of the covariate, missing the potential impact of the preceding covariate history of smoking risk. A model describing the impact of the preceding covariate history might have intensity taking the form $\lambda(t) = \exp \left\{ \int_{s < t} f(z(t-s); \theta) ds \right\}$, where $\exp \{f(z(t-s); \theta)\}$ typically would be a decreasing function of time. In this case, the cumulative intensity would involve a double integral of the form $\int_A \exp \left\{ \int_{s < t} f(z(t-s); \theta) ds \right\} d\mathbf{z}$, for which design-unbiased estimators may be constructed. Our current research involves fixed effects models for this case, but we have yet to considered mixed effects versions.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH grant 1RO1DA024687. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Institutes of Health. We also thank the Associate Editor and the referee whose comments improved the quality of our paper.

References

- Anderson, TW. An Introduction to Multivariate Statistical Analysis. New York: Wiley; 1984.
- Bellman, RE. Dynamic Programming. Princeton, N.J.: Princeton University Press; 1957.
- Booth JG, Holbert JP. Maximizing generalized linear mixed model likelihood with an automated Monte Carlo EM algorithm. *Journal of the Royal Statistical Society, Series B*. 1999; 61:265–285.
- Breslow NE, Clayton DG. Approximate inference in Generalized Linear Mixed Models. *Journal of the American Statistical Association*. 1993; 88:9–25.
- Breslow NE, Lin X. Bias correction in generalized linear mixed models with a single component of dispersion. *Biometrika*. 1995; 82:81–91.
- Breslow NE, Lin X. Analysis of correlated binomial data in logistic-normal models. *Journal of Statistical Computation and Simulation*. 1996; 55:133–146.
- Brewer, MB. Research design and issues of validity. In: Reis, H.; Judd, C., editors. *Handbook of Research Methods in Social and Personality Psychology*. Cambridge: Cambridge University Press; 2000.
- Cassel, C-M.; Särndal, C-E.; Wretman, JH. Foundations of inference in survey sampling. New York: Wiley; 1977.
- Cole DJ, Morgan BJT, Ridout MS. Generalized linear mixed models for strawberry inflorescence data. *Statistical Modelling*. 2003; 3:273–290.
- Collins RL, Morsheimer ET, Shiffman S, Paty JA, Gnys M, Papandonatos GD. Ecological momentary assessment in a behavioral drinking moderation training program. *Experimental and Clinical Psychopharmacology*. 1998; 6:306–315. [PubMed: 9725114]
- Daley, DJ.; Vere-Jones, D. An Introduction to the Theory of Point Processes. New York: Springer-Verlag; 1988.
- Ferreira PE. Multiparametric estimating equations. *Annals of the Institute for Mathematical Statistics*. 1982; 34A:423–431.
- Freedman MJ, Lester KM, McNamara C, Milby JB, Schumacher JE. Cell phones for ecological momentary assessment with cocaine-addicted homeless patients in treatment. *Journal of Substance Abuse Treatment*. 2006; 30:105–111. [PubMed: 16490673]
- Gill RD, Johansen S. A survey of product-integration with a view toward application in survival analysis. *Annals of Statistics*. 1990; 18:1501–1555.
- Ha ID, Lee Y. Estimating frailty models via Poisson hierarchical generalized linear models. *Journal of Computational and Graphical Statistics*. 2003; 12:663–681.
- Ha ID, Lee Y, Song J-K. Hierarchical likelihood approach to frailty models. *Biometrika*. 2001; 88:233–243.
- Hufford MR, Shiffman S. Assessment methods for patient-reported outcomes. *Disease Management and Health Outcomes*. 2003; 11:77–86.
- Kassel JD, Stroud LR, Paronis CA. Smoking, stress, and negative affect: Correlation, causation, and context across stages of smoking. *Psychological Bulletin*. 2003; 129:270–304. [PubMed: 12696841]
- Kuk AYC. Laplace importance sampling for generalized linear mixed models. *Journal of Statistical Computation and Simulation*. 1999; 63:143–158.

- Lang K, Hunter DR, Yang I. Optimization transfer using surrogate objective functions. *Journal of Computational and Graphical Statistics*. 2000; 9:1–20.
- Lawless JF. Regression methods for Poisson process data. *Journal of the American Statistical Association*. 1987; 82:808–815.
- Lee Y, Nelder JA. Hierarchical generalized linear models. *Journal of the Royal Statistical Society, B*. 1996; 58:619–678.
- Liu L, Huang XL. Joint analysis of correlated repeated measures and recurrent events processes in the presence of death with application to a study on acquired immune deficiency syndrome. *Applied Statistics*. 2007; 58:65–81.
- McCulloch CE. Maximum likelihood algorithms for Generalized Linear Mixed Models. *Journal of the American Statistical Association*. 1997; 92:162–170.
- Meng X-L, Rubin DB. Using EM to obtain asymptotic variance-covariance matrices: The SEM algorithm. *Journal of the American Statistical Association*. 1991; 86:899–909.
- Neustifter B, Rathbun SL, Shiffman S. Mixed-Poisson point process with partially-observed covariates: Ecological Momentary Assessment of smoking. *Journal of Applied Statistics*. 2012; 39:883–899. [PubMed: 22544991]
- Pinheiro JC, Bates DM. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics*. 1995; 4:12–35.
- Rathbun SL. Optimal estimation of Poisson intensity with partially observed covariates. *Biometrika*. 2013; 100:277–281.
- Rathbun SL, Shiffman S, Gwaltney CJ. Modelling the effects of partially observed covariates on Poisson process intensity. *Biometrika*. 2007; 94:153–165.
- Rathbun SL, Song X, Neustifter B, Shiffman S. Survival analysis with time varying covariates measured at random times by design. *Applied Statistics*. 2013; 62:419–434. [PubMed: 23729866]
- Rizopoulos D, Verbeke G, Lesaffre E. Fully exponential Laplace approximations for the joint modelling of survival and longitudinal data. *Journal of the Royal Statistical Society, Series B*. 2009; 71:637–654.
- Robinson MD, Clore GL. Belief and feeling: Evidence for an accessibility model of emotional self-report. *Psychological Bulletin*. 2002; 128:934–960. [PubMed: 12405138]
- Russell J. A circumplex model of affect. *Journal of Personality and Social Psychology*. 1980; 37:345–56.
- Sayette M, Hufford M. Effects of cue exposure and deprivation on cognitive resources in smokers. *Journal of Abnormal Psychology*. 1994; 103:812–818. [PubMed: 7822584]
- Schadel WG, Shiffman S, Niaura R, Nichter M, Abrams DB. Current models of nicotine dependence: What is known and what is needed to advance understanding of tobacco etiology among youth. *Drug and Alcohol Dependence*. 2000; 59:S9–S22. [PubMed: 10773435]
- Shiffman S. Dynamic influences on smoking relapse process. *Journal of Personality*. 2005; 73:1715–1748. [PubMed: 16274451]
- Shiffman S, Stone AA. Introduction to the special section: Ecological momentary assessment in health psychology. *Health Psychology*. 1998; 17:3–5.
- Shiffman S, Waters AJ. Negative affect and smoking lapses: A prospective analysis. *Journal of Consulting and Clinical Psychology*. 2004; 72:192–201. [PubMed: 15065954]
- Shiffman S, Paty JA, Gnys M, Kassel JD, Hickcox M. First lapses to smoking: Within subjects analysis of real time reports. *Journal of Consulting and Clinical Psychology*. 1996; 64:366–379.
- Shiffman S, Gwaltney CJ, Baladonis MH, Liu KS, Paty JA, Kassel JD, et al. Immediate antecedents of cigarette smoking: An analysis from ecological momentary assessment. *Journal of Abnormal Psychology*. 2002; 111:531–545. [PubMed: 12428767]
- Shiffman S, Baladonis MH, Gwaltney CJ, Paty JA, Gnys M, Kassel JD, et al. Prediction of lapse from associations between smoking and situational antecedents assessed by ecological momentary assessment. *Drug and Alcohol Dependence*. 2007; 91:159–168. [PubMed: 17628353]
- Slavík, A. *Product Integration, its History and Applications*. Prague: Matfyzpress; 2007.
- Steele BM. A modified EM algorithm for estimation in generalized mixed models. *Biometrics*. 1996; 52:1295–1310. [PubMed: 8962454]

- Stone AA, Shiffman S. Ecological momentary assessment (EMA) in behavioral medicine. *Annals of Behavioral Medicine*. 1994; 16:199–202/.
- Stone, AA.; Shiffman, S.; Atienza, AA.; Nebeling, L. *Self Reports in Public Health*. New York: Oxford University Press; 2007. *The Science of Real- Time Data Capture*.
- Strecher VJ. Computer-tailored smoking cessation materials: A review and discussion. *Patient Education and Counseling*. 1999; 36:107–117. [PubMed: 10223016]
- Thall PF. Mixed Poisson likelihood regression models for longitudinal interval count data. *Biometrics*. 1988; 44:197–209. [PubMed: 3358988]
- Tierney L, Kass RE, Kadane JB. Fully exponential Laplace approximations for expectations and variances of nonpositive functions. *Journal of the American Statistical Association*. 1989; 84:710–716.
- Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: an overview. *Statistica Sinica*. 2004; 14:809–834.
- United States Department of Health and Human Services. *The Health Consequences of Smoking—50 Years of Progress. A Report of the Surgeon General*. Atlanta, Georgia: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health; 2014.
- Waagepetersen R. Estimating functions for inhomogeneous spatial point processes with incomplete covariate data. *Biometrika*. 2008; 95:351–363.
- Warren CA, McDonough BE. Event-related brain potentials as indicators of smoking cue-reactivity. *Clinical Neurophysiology*. 1999; 110:1570–1784. [PubMed: 10479024]
- Wulfsohn MS, Tsiatis AA. A joint model for survival and longitudinal data measured with error. *Biometrics*. 1997; 53:330–339. [PubMed: 9147598]
- Zhang HP, Ye YQ, Diggle P, Shi J. Joint modeling of time series measures and recurrent events and analysis of the effects of air quality on respiratory systems. *Journal of the American Statistical Association*. 2008; 103:48–60.

Figure 1.
Distributions of Random Covariate Effects.

Table 1

Parameter estimates for fixed effects, frailty and mixed-effects models. The last column gives the standard deviation of the random effects.

Covariate	Fixed Effects			Frailty*			Mixed Effect		
	Estimate	SE	p-value	SE	p-value	Estimate	SE	p-value	SD
Intercept	-0.008	0.021		0.074		-0.103	0.107		0.509
Negative Affect	0.046	0.009	<0.0001	0.027	0.0914	0.096	0.025	0.0002	0.106
Restlessness	0.082	0.007	<0.0001	0.031	0.0072	0.256	0.065	<0.0001	0.226
Alone	0.155	0.018	<0.0001	0.039	0.0001	0.175	0.053	0.0010	0.394
Other Smokers	0.263	0.021	<0.0001	0.062	<0.0001	0.434	0.082	<0.0001	0.456
Prohibited	-0.952	0.033	<0.0001	0.091	<0.0001	-0.904	0.072	<0.0001	
Allowed	0.309	0.025	<0.0001	0.084	0.0002	0.402	0.068	<0.0001	

* Parameter estimates for frailty model are equal to those of the fixed effects model.

Table 2

Estimated risk ratios for the effects of mood variables on smoking rate. Smoking rate in cigarettes per hour are given for the polychotomous environment variables.

Covariate	Risk ratios and smoking rates	95% confidence Interval
Negative Affect *	1.104	1.051, 1.160
Restlessness *	1.305	1.158, 1.471
Not Alone [†]	0.723	0.602, 0.869
Alone [†]	0.866	0.708, 1.060
No Other Smokers Present [†]	0.742	0.616, 0.895
Other Smokers Present [†]	1.166	0.941, 1.445
Prohibited [†]	0.371	0.307, 0.449
Discouraged [†]	0.944	0.769, 1.158
Allowed [†]	1.424	1.193, 1.701

* Estimated risk ratios.

[†] Estimated smoking rates in cigarettes per hour.

Table 3

Correlations among random effects.

	Intercept	Negative Affect	Restlessness	Alone	Other Smokers
Intercept	1.000				
Negative Affect	0.224	1.000			
Restlessness	-0.138	0.723	1.000		
Alone	-0.256	-0.106	-0.046	1.000	
Other Smokers	-0.426	-0.010	0.027	0.467	1.000