2018

A Mixed-effects Location Scale Model for Time-to-event Data: A Smoking Behavior Application

Delphine Courvoisier

Theodore A. Walls
University of Rhode Island, walls@uri.edu

See next page for additional authors

Citation/Publisher Attribution
Courvoisier, D., Walls, T. A., Cheval, B., & Hedeker, D. (2018). A Mixed-effects Location Scale Model for Time-to-event Data: A Smoking Behavior Application. *Addictive Behaviors*. In press.
Available at: https://doi.org/10.1016/j.addbeh.2018.08.032

A Mixed-effects Location Scale Model for Time-to-event Data: A Smoking Behavior Application

Delphine Courvoisier[1]
University of Geneva and University Hospitals of Geneva[1]
Theodore A. Walls
University of Rhode Island
Boris Cheval[1]
University of Geneva and University Hospitals of Geneva[1]
Donald Hedeker
University of Chicago

Author Note:

Delphine Courvoisier, Faculty of Medicine, University of Geneva, Switzerland

Boris Cheval, Faculty of Medicine, University of Geneva, Switzerland

Theodore A. Walls, Department of Psychology, University of Rhode Island

Donald Hedeker, Department of Biostatistics, University of Chicago

Correspondence concerning this article should be addressed to: Delphine Courvoisier,

Delphine.Courvoisier@hcuge.ch

Paper accepted for publication in Addictive behaviors (August 2018)

Abstract

In general, mixed-effects location scale models (MELS) allow assessment of within-person and between-person variability with time-to-event data for outcomes that follow a normal or ordinal distribution. In this article, we extend the mixed-effects location scale model to time-to-event data in relation to smoking data. Better understanding of the time-graded within-person variability of factors involved in nicotine dependence can be helpful to researchers in their efforts to fine-tune smoking cessation programs. We illustrate the MELS model with data on time to first cigarette measured every day for 7 days in smokers randomized to two groups: a) those asked to keep smoking, or b) those asked to stop. Our results show that some individuals remain very stable in their time to first cigarette over the week, while others show variable patterns. The stable individuals smoked every day, did not smoke immediately upon waking, and were all in the group asked to keep smoking. Conversely, the variable individuals had at least one day during which they did not smoke, other days during which they smoked within the first 5 minutes of waking, and they were almost all in the group asked to quit smoking. These findings suggested that MELS have the potential to provide insights on how people try to stop smoking. More importantly, this model can be applied to other clinically important outcomes such as time to relapse in a range of cessation programs.

Keywords: Longitudinal data; random effects; mixed-effects location scale models; time to first cigarette

Highlights: MELS for time to event data are an extension of mixed effect models.

MELS can answer questions on factors of between- and within-person variability.

Time to first cigarette was more variable among 851 smokers asked to quit.

Time to first cigarette was more variable from day to day among smokers asked to quit

Role of funding sources: This project was supported by the Swiss National Fund (PASMP3_136980). Dr. Walls was supported by and award from the American Cancer Society MSRG 07-015-01.

Declarations of interest: none

Contributors: All authors contributed to the manuscript. Drs. Courvoisier and Walls framed the project and developed it. Dr. Hedeker provided assistance with the software and advice on the approach. Dr. Cheval participated in drafting the manuscript.

A Mixed-effects Location Scale Model for Time-to-event Data: A Smoking Behavior Application

Recovery from nicotine dependence requires addressing many factors which can detract from individuals' efforts to achieve abstinence. A primary indicator of nicotine dependence is the time that elapses until the first cigarette after waking (Baker et al., 2007; Branstetter & Muscat, 2013; Fu et al., 2012; Haberstick et al., 2007; Heatherton, Kozlowski, Frecker, Rickert, & Robinson, 1989; Muscat, Ahn, Richie, & Stellman, 2011; Muscat, Liu, Livelsberger, Richie, & Stellman, 2012; Muscat, Stellman, Caraballo, & Richie, 2009). Previous studies examining the time to first cigarette have generally focused on between individual differences likely to impact this time, but rarely on factors that may vary within each individual (Haller, Etter, & Courvoisier, 2014). In this case, understanding why some individuals are less able to delay their first cigarette use than others (i.e., between-subject variability), as well as under what conditions a given individual is more or less likely to delay this first use (i.e., within-subject variability) are both of equal clinical importance. However, statistical assessment of time-to-event effects on an ordinal outcome when the within and between person variability is of interest requires specialized statistical modeling.

Mixed-effects location scale models (MELS) have been developed to allow estimation of these two sources of variability. That is, MELS allows researchers to include covariates not only on the random (between-subject) intercept, but also on the random (within-subject) residual variance (Hedeker, Mermelstein, & Demirtas, 2008). By including covariates on the **between-subject** variance, researchers can utilize this model component to answer questions such as whether smoking cessation program A has more similar (less variable) impact over all individuals (first time to cigarette was, on average, 50 minutes ± 5 minutes) than smoking cessation program B (first time to cigarette was, on average, 50 minutes ± 40 minutes). At the same time, including covariates on the **within-subject** variance allows researchers to test

hypotheses about how well specific measurements can be estimated within each individual. This model component handles questions like: are some individuals in the sample more variable than other individuals or, conversely, are some more predictable than others? For instance, smokers of more than 10 years may be more predictable over time than more recent smokers.

These powerful models were developed only for outcomes that follow a normal (Hedeker et al., 2008; Hedeker, Mermelstein, & Demirtas, 2012), or ordinal (Hedeker, Demirtas, & Mermelstein, 2008) distribution, and not for time-to-event response distributions. In this article, we extend the MELS model to time-to-event data, as such outcomes are of particular importance in addiction. We first present the MELS model. Then, we describe how to extend the classic mixed-effects location model (i.e., only random location) to time-to-event data by considering these data as a form of ordinal data. Finally, we develop the MELS model for time-to-event data. We illustrate the model by applying it to data on time to first cigarette, measured every day for 7 days in ongoing and just stopping smokers. Thus, some people will not have the event (right censored), while others will still smoke. One way of considering time to event is to separate the time into discrete intervals. This strategy has several advantages. The first advantage is that the intervals can be chosen to be of equal sizes (i.e., each bin is 5 minutes), but can also be of unequal sizes to better reflect meaningful distinctions. For instance, the time before smoking the first cigarette of the day is an indicator of cigarette dependence. Smoking within the first five minutes after waking up indicates a strong dependence, whereas the distinction between 30 and 35 minutes after waking up is of lesser importance. A second advantage is that the proportional hazard assumption can easily be tested by allowing the baseline odds to vary. A third advantage is that continuous time model often has difficulties with ties. These ties are frequent when time to event is self-reported because respondents often round the exact time (e.g., time to first cigarette). A fourth

advantage is that right censoring can easily be accommodated by adding a final interval, corresponding to 'no occurrence of the event' (e.g., not smoking the whole day in the case of time to first cigarette). Finally, a fifth, practical, advantage is that almost no precision is lost by using a discrete approach, and there is more available software for these analyses.(Liu & Huang, 2008)

Mixed-effects Location Scale Models

Mixed effects location scale models for Gaussian response (Hedeker et al., 2008, 2012) use the following model:

$$\boldsymbol{\gamma_{ij}} = \boldsymbol{x'_{ij}\beta} + \boldsymbol{v_i} + \boldsymbol{\epsilon_{ij}} \ (\mathbf{1})$$

Where $\boldsymbol{\gamma_{ij}}$ is the outcome variable for individual $\boldsymbol{i}$ ($\boldsymbol{i} = \mathbf{1}, \mathbf{2}, \ldots, \boldsymbol{N}$) at occasion $\boldsymbol{j}$ ($\boldsymbol{j} = \mathbf{1}, \mathbf{2}, \ldots, \boldsymbol{n_i}$ occasions), $\boldsymbol{x_{ij}}$ is a vector of $\boldsymbol{p}$ predictors, usually including 1 for the intercept, and $\boldsymbol{\beta}$ is the corresponding vector of $\boldsymbol{p}$ regression coefficients. $\boldsymbol{v_i}$ is a random intercept and can be assumed to follow a normal distribution with mean zero and variance $\boldsymbol{\sigma_v^2}$, representing the between-subject (BS) variance. The residuals $\boldsymbol{\epsilon_{ij}}$ are also assumed to follow a normal distribution with mean zero and variance $\boldsymbol{\sigma_\epsilon^2}$, independent of the random effects. This variance represents the within-subject (WS) variance.

With respect to the random effects, it is possible to include covariates on the BS and on the WS variance, and also to allow the WS variance of the residuals to vary by individual ($\boldsymbol{\omega_i}$). One possible way to do this is to use a log-linear representation:

$$\boldsymbol{\sigma^2_{v_{ij}}} = \mathbf{exp}(\boldsymbol{u'_{ij}\alpha}) \ (\mathbf{2})$$

$$\boldsymbol{\sigma^2_{\epsilon ij}} = \mathbf{exp}(\boldsymbol{\omega'_{ij}\tau} + \boldsymbol{\omega_i}) \ (\mathbf{3})$$

where $\boldsymbol{u_{ij}}$ and $\boldsymbol{\omega_{ij}}$ are covariates on the BS and WS variance respectively, and $\boldsymbol{\alpha}$ and $\boldsymbol{\tau}$ are the coefficients of the covariates for the BS and WS variance respectively. The random location scale effect $\boldsymbol{\omega_i}$ follow a normal distribution with mean zero and variance $\boldsymbol{\sigma_\omega^2}$. Note that the variances now have subscripts $\boldsymbol{i}$ and $\boldsymbol{j}$ to indicate that their values change based on the values

and coefficients of the covariates. Since $\boldsymbol{\omega_i}$ is specified as following a normal distribution, the distribution of $\boldsymbol{\sigma^2_{\epsilon ij}}$ follows a log-normal distribution, and is thus always positive and skewed to the right, making it a good choice for this variance parameter. Since MELS' parameters are estimated using full likelihood estimation, it provides valid inference for incomplete data under missing at random (MAR) assumptions (Little and Rubin, 2002).

This model has already been extended to ordinal outcomes (Hedeker et al., 2009). Because time to event data can be thought of corresponding to an ordinal scale, with each period of time potentially realizing (or not) the event, it is possible to extend the random location scale model to time to event data within the existing multilevel modeling structures for ordinal data.

Mixed-effects location model for time to event data using an ordinal specification

An ordinal specification of a mixed effects model (with only random location) can be used to model time-to-event data by dividing the time into several time intervals (Grilli, 2005; Hedeker & Mermelstein, 2011; Hedeker, Siddiqui, & Hu, 2000). An additional category can be used to allow for censoring. Let $\boldsymbol{i}$ be the level-2 units ($\boldsymbol{i = 1, \ldots, N}$) and $\boldsymbol{j}$ be the level-1 units ($\boldsymbol{j = 1, 2, \ldots, n_i}$). In the case of repeated events per subject, the level-2 units are the subjects and the level-1 units represent the repeated event times. Even assuming a continuous random variable for the uncensored time of event occurrence, time of assessment often occurs at discrete positive values $\boldsymbol{t = 1, 2, \ldots, m}$. For each level-1 unit, observations continue until time $\boldsymbol{t_{ij}}$, when the time index for estimation ends. The additional category used for censoring will be used for those observations that still have not experienced the event at time $\boldsymbol{m}$. The probability of event occurrence up to and including time interval $\boldsymbol{t}$, $\boldsymbol{P_{ijt}}$ is

$$\boldsymbol{P_{ijt} = Pr\left[t_{ij} \leq t\right], (4)}$$

and the survivor function - the probability of the event not occurring before or at time $\boldsymbol{t}$ - is the complement $\boldsymbol{1 - P_{ijt}}$.

Ordinal data can be modeled through the logit, the probit and the complementary log-log link. In this presentation, we will focus on the logit link. Predicting the survivor function with a linear regression model using a logit link leads to the following equation:

$$\mathbf{log}\left[\boldsymbol{P_{ijt}}/(\mathbf{1}-\boldsymbol{P_{ijt}})\right] = \boldsymbol{\gamma_t} + \left(\boldsymbol{x'_{ij}\beta} + \boldsymbol{v_i}\right), (\mathbf{5})$$

or, equivalently,

$$\boldsymbol{P_{ijt}} = \mathbf{1}/(\mathbf{1} + \mathbf{exp}(-(\boldsymbol{\gamma_t} + \boldsymbol{x'_{ij}\beta} + \boldsymbol{v_i}))) , (\mathbf{6})$$

where $\boldsymbol{\gamma_t}$ is a set of $\boldsymbol{t}-\mathbf{1}$ strictly increasing thresholds, and $\boldsymbol{x_{ij}}$ is a vector of size $\boldsymbol{p}$ including covariates. These covariates may vary at the subject (level-2) or at the repeated events level (level-1). However, they may not vary across (ordinal) time points $\boldsymbol{t}$. Thus, for instance, covariates could be gender (level-2), day of the week (level-1), but not number of persons present at time $\boldsymbol{t}$, which varies at each assessment of event occurrence. If covariates at time $\boldsymbol{t}$ are of interest, they can be included as an average over the whole time or given as the specific value of the covariate at the time of event occurrence. $\boldsymbol{v_i}$ is a random intercept at the level-2 unit $\boldsymbol{i}$ and is assumed to follow a normal distribution of mean 0 and variance $\boldsymbol{\sigma^2_{v_i}}$. Note that this leads to the estimation of odds ratios.

Alternatively, each survival time can be represented as a $\boldsymbol{t_{ij}} \times \mathbf{1}$ vector of zeros for censored individuals, and a $(\boldsymbol{t_{ij}}-\mathbf{1}) \times \mathbf{1}$ vector of zeros followed by a 1 for individuals experiencing the event. These indicators can then be considered as the outcome of a multilevel dichotomous regression model.(Reardon, Brennan, & Buka, 2002; Singer & Willett, 2003)

The probability of event occurrence in time interval $[t_{k-1}, t_k]$, conditional on non-occurrence prior to $[t_{k-1}]$, can be defined as p_{ijt}:

$$p_{ijt} = Pr\left[\mathrm{t_{ij}} \in [t_{k-1}, t_k] \mid \mathrm{t_{ij}} > t_{k-1}\right]. (7)$$

Thus, the complement $1 - p_{ijt}$ is the probability of the event not occurring beyond time

interval $[t_{k-1}, t_k]$, conditional on it not occurring prior to t_{k-1}. Including random effects, the proportional odds model can then be written as:

$$\log\left[p_{ijt}/(1-p_{ijt})\right] = \gamma_t + \left(x'_{ijt}\beta + v_i\right). (8)$$

where v_i is assumed to follow a normal distribution of mean 0 and variance $\sigma^2_{v_i}$.

Compared to the ordinal specification, the dichotomous approach is necessary to include covariates changing within each time interval (e.g., number of people present within each time interval). It also allows the possibility of relaxing the proportional hazards assumption (e.g., the impact of gender differs when considering the transition between various time intervals). The main disadvantage of the dichotomous approach is the size of the data matrix. In the ordinal approach, the outcome consists of two pieces of information: the (ordinal) time of the event and whether or not it was censored. In the case of an outcome where censoring can only occur at the end of the observation period with no event, these two pieces of information can be presented more simply by adding one interval $m + 1$ to the time interval variable. However, in the dichotomous approach, each survival time is represented as a vector of indicators, and the length of the vector depends on the censoring, the timing of the event, and the number of prespecified time intervals.

Mixed effects location scale model for time to event data

The random location scale model for ordinal data, treating time as a series of dichotomous variables, including only a random intercept and using a logit link (Hedeker et al., 2009), uses equation 8 for the expected value but the between-subject and within-subject variances can vary according to covariates u and ω respectively. Furthermore, the WS-variance can also vary across individuals (ω_i).

$$\sigma^2_{v_i} = \exp(u'_i\alpha) (9)$$

$$\sigma^2_{\epsilon_{ijt}} = \exp\left(\omega'_{ij}\tau + \omega_i\right) (10)$$

By including a first column of one in u'_i, the reference between-subject variance becomes

$\exp(\alpha_0)$ (see Appendix A for details on the estimation)

Material and methods

Smokers coming to a stop smoking website were invited to participate in a study on symptoms' evolution the first few days after quitting smoking. After a baseline assessment at Day 0, participants were randomized (1:1) into one group asked to stop smoking, and one group asked to keep smoking as usual. Their smoking behavior and their symptoms were then assessed every day for the next 7 days. To be included in the study, participants had to be daily smokers, older than 18 years old, and willing to fill in a short online questionnaire per day over one week. Participants were excluded if they answered less than 2 questionnaires. Participants signed an informed consent before filling out the baseline questionnaire. In this example, we focus on time to first cigarette as the outcome, and group, day of measurement, and their interaction as the predictors.

Smoking behavior indicators

Three questions were employed to reflect smoking behavior.

Smoking. Smoking was assessed with the item "Do you currently smoke tobacco or cannabis?" Participants had to choose one of the following options: yes, I smoke every day; yes occasionally (NOT every day); no I quit smoking; no I have never been a smoker. Participants who answered "yes, I smoke every day" were classified as daily smokers.

Number of cigarette per day. The habitual number of cigarette per day was assessed with the item "On average, how many cigarettes do you smoke per day, currently?" Participants had to select a number from 0 to 100 cigarettes per day.

Time to first cigarette. The habitual time to first cigarette was assessed with the item "Usually, how soon after you wake up do you smoke your first cigarette of the day?" Participants had to select a number from less than 1 minutes to 16 hours. Time to first cigarette was then categorized into 5 categories: 1-5 minutes; 6-15 minutes; 16-30 minutes;

31-60 minutes and >60 minutes (Etter, Le Houezec, & Perneger, 2003). A sixth category was added to include not smoking at all during the day.

Statistical analyses

Data were analyzed by using the mixed effects random location model to predict the time to first cigarette using the PROC NLMIXED procedure (see SAS code, appendix 2). This procedure allows estimating a broad range of models due to its flexibility and set of estimation techniques. It is currently difficult to estimate these models in R, or other statistical software, without some programming. We first estimated a mixed effects random location model with the three main effects (group, day, group*day) and only a single random intercept. The reference group was patients asked to keep smoking. Then, using the same fixed effects, we included group as a predictor of both between- and within-subject variability. Starting values for the MELS model were chosen using the initial mixed effects random location model.

Results

The total sample included 851 smokers (Table 1). The two randomized groups were similar in gender distribution, age, number of cigarette per day, and time to first cigarette (Table 2). A classic mixed effect model with the three main effects (group, day, group*day) and only a single random intercept showed, as expected, that smoking a cigarette occurs significantly later for smokers asked to stop smoking (Table 3, left columns). Across the days of study, smokers asked to keep smoking smoked their first cigarette earlier and earlier (effect of day: p=0.001). The interaction effect showed that time to first cigarette decreased similarly for smokers asked to quit smoking (p=0.14).

Using the same fixed effects, but including group as a predictor of both between- and within-subject variability, the effect of group had a relatively similar estimate in both models, but the decrease across day in time to first cigarette was much smaller and non-significant.

Note that a change in the stochastic structure usually leads to reduced bias in the standard errors, but not in the coefficients themselves. The change in coefficients in this application may be due to the improved model specification or the more appropriate accounting of measurement error at the between- and within-individual level. The between-subject variability was larger for smokers asked to keep smoking than for those asked to stop. Within-subject variability indicated that smokers asked to keep smoking were significantly more predictable (i.e., less variable) than smokers asked to quit.

At the individual level, the random scale parameter estimated by a MELS model with group, day, and their interaction as fixed effects and group as a predictor of between-subject variance, showed a relatively large spread (supplementary figure 1), compared to the between-subject variance. In this model, we did not include group as a predictor of within-subject variance to order to assess how being asked to quit smoking, compared to being asked to continue smoking, changed the predictability of smoking patterns. These parameters showed that some individuals remain very stable in their time to first cigarette over the week, while others showed very variable patterns (for a boxplot of all individual random scale estimation, see supplementary figure 1). As an illustration, Figure 1 (Panel A) shows time to first cigarette in categories over time for the 10 individuals with the lowest estimates of the random scale (ω_i), and Figure 1 (Panel B) shows the same trajectories but for the 10 individuals with the highest estimates of the random scale. The ten stable individuals smoked every day and did not smoke immediately upon waking. They were all in the group asked to keep smoking. In contrast, most of the ten variable individuals had at least one day during which they did not smoke, and other days during which they smoked within the first 5 minutes of waking. Nine of the ten variable individuals were in the group asked to quit smoking.

A model allowing the baseline odds to vary did not have a better fit than the presented

model, thus supporting the use of a proportional hazard model.

Discussion

Because of its clinical importance, a better understanding of the factors influencing the first time to cigarette, an accurate indicator of nicotine dependence (Baker et al., 2007; Branstetter & Muscat, 2013; Fu et al., 2012; Muscat et al., 2011; Muscat et al., 2012; Muscat et al., 2009), is important. However, previous studies have generally focused on individual differences likely to impact this first time to cigarette, but rarely on factors that may vary within each individual. As such outcome is of particular interest in addiction, in this article, we have modified an existing mixed effects location scale model proposed by Hedeker and colleagues (2008) to apply to time to event data, allowing for right censoring. This modification was done by dividing time to events in intervals, and using a proportional odds specification of the mixed effects location scale model. The use of a final interval to account for right censoring (i.e., not smoking the whole day when time to event is time to first cigarette) is particularly appropriate in the case of smoking patterns. However, it could be less optimal for other outcomes such as time to relapse, where the event would have to occur in the framework of the MELS models including repeated measurements over time. We also discussed other specifications, for instance with a complementary log-log link, that would yield other estimators (i.e., hazard ratio with the complementary log-log link). The MELS model allows adding covariates on the variances at the subject and day level, and also permits a variability in how predictable each individual is by including a subject-level random scale effect on the within-subject variance. A practical advantage of this model is that it can easily be carried out using existing software (SAS PROC NLMIXED).

Our simple example with time to first cigarette, measured each day for seven days, and only one predictor, shows that the MELS model has a much better fit, as assessed by information criteria, than a classic linear mixed effect model with only a random intercept and

slope. Furthermore, fixed effects can be significantly modified by a more correct specification of the random structure, maybe due to a better model specification or a more appropriate accounting of measurement error. However, we cannot exclude that the change in coefficients is a consequence of the estimation method of PROC NLMIXED (Adaptive Gaussian quadrature), which can be sensitive to starting values (Wang & McArdle, 2008). In this analysis, we followed the suggestion of Li and Hedeker (2002) and chose starting values based on the initial classic mixed effect models, as random starting values resulted in model convergence issues. Finally, results revealed that some individuals remain very stable in their time to first cigarette over the week, while others show very variable patterns. Interestingly, the stable individuals smoked every day, did not smoke immediately upon waking, and were all in the group asked to keep smoking. Conversely, the variable individuals had at least one day during which they did not smoke, other days during which they smoked within the first 5 minutes of waking, and were almost all in the group asked to quit smoking. However, it should be noted that the loss to follow-up of the sample during the 7 days follow-up period is extensive (58%) and that a substantial portion of the participants did not met the inclusion criterion of at least two measurements (18%). As such, a selection bias due to attrition cannot be excluded. In addition, the present illustration did not include time-varying predictors (e.g., self-efficacy, motivation to quit) or non-time-varying predictors (e.g., longest duration of previous abstinence from tobacco, nicotine dependence). Overall, these findings highlighted that MELS have the potential to provide insights on how people try to stop smoking. The model can be used for other clinically important outcomes such as time to relapse in cessation programs.

Several extensions of the initial MELS model have now been developed (Hedeker & Nordgren, 2013), including MELS for ordinal data (Hedeker, et al., 2008), for 3-level continuous data, (Li & Hedeker, 2012), and for bivariate outcomes (Pugach, Hedeker,

Richmond, Sokolovsky, & Mermelstein, 2014). Given the usefulness of this model, further research may be warranted to provide a comprehensive framework of the mixed-effects location scale model for diverse types of outcomes and various random structures.

References

Baker, T. B., Piper, M. E., McCarthy, D. E., Bolt, D. M., Smith, S. S., Kim, S.-Y., . . . Toll, B. A. (2007). Time to First Cigarette in the Morning as an Index of Ability to Quit Smoking: Implications for Nicotine Dependence. *Nicotine & Tobacco Research,* 9, S555-S570. doi: 10.1080/14622200701673480

Branstetter, S. A., & Muscat, J. E. (2013). Time to first cigarette and 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanol (NNAL) levels in adult smokers; National Health and Nutrition Examination Survey (NHANES), 2007–2010. *Cancer Epidemiology and Prevention Biomarkers,* 22, 615-622. doi: 10.1158/1055-9965

Etter, J. F., Le Houezec, J., & Perneger, T. V. (2003). A self-administered questionnaire to measure dependence on cigarettes: the cigarette dependence scale. *Neuropsychopharmacology*, *28*, 359-370.

Fu, M., Martínez-Sánchez, J. M., Agudo, A., Pascual, J. A., Ariza, C., Moncada, A., . . . Investigators, D. S. (2012). Nicotine depedence and salivary cotinine concentration in daily smokers. *European Journal of Cancer Prevention,* 21, 96-102. doi: 10.1097/CEJ.0b013e32834a7e59

Grilli, L. (2005). The random-effects proportional hazards model with grouped survival data: a comparison between the grouped continuous and continuation ratio versions. *Journal of the Royal Statistical Society Series a-Statistics in Society,* 168, 83-94. doi: 10.1111/j.1467-985X.2004.00337.x

Haberstick, B. C., Timberlake, D., Ehringer, M. A., Lessem, J. M., Hopfer, C. J., Smolen, A., & Hewitt, J. K. (2007). Genes, time to first cigarette and nicotine dependence in a general population sample of young adults. *Addiction,* 102, 655-665. doi: 10.1111/j.1360-0443.2007.01746.x

Haller, C. S., Etter, J.-F., & Courvoisier, D. S. (2014). Trajectories in cigarette dependence as a function of anxiety: a multilevel analysis. *Drug and alcohol dependence,* 139, 115-120. doi: 10.1016/j.drugalcdep

Heatherton, T. F., Kozlowski, L. T., Frecker, R. C., Rickert, W., & Robinson, J. (1989). Measuring the heaviness of smoking: using self-reported time to the first cigarette of the day and number of cigarettes smoked per day. *Addiction,* 84, 791-800. doi: 10.1016/j.drugalcdep

Hedeker, D., Demirtas, H., & Mermelstein, R. J. (2009). A mixed ordinal location scale model for analysis of Ecological Momentary Assessment (EMA) data. *Statistics and Its Interface,* 2, 391-401.

Hedeker, D., & Mermelstein, R. J. (2011). *Multilevel analysis of ordinal outcomes related to survival data*. In J.J. Hox & J. K. Roberts (Eds.), *Handbook of advanced multilevel analysis.* (pp. 115-136). New York, NY: Routledge

Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2008). An application of a mixed-effects location scale model for analysis of Ecological Momentary Assessment (EMA) data. *Biometrics,* 64, 627-634. doi: 10.1111/j.1541-0420.2007.00924.x

Hedeker, D., Mermelstein, R. J., & Demirtas, H. (2012). Modeling between-subject and within-subject variances in ecological momentary assessment data using mixed-effects location scale models. *Statistics in Medicine,* 31, 3328-3336. doi: 10.1002/sim.5338

Hedeker, D., & Nordgren, R. (2013). MIXREGLS: A Program for mixed-effects location scale analysis. *Journal of Statistical Software,* 52, 1-38. doi: 10.18637/jss.v052.i12

Hedeker, D., Siddiqui, O., & Hu, F. B. (2000). Random-effects regression analysis of correlated grouped-time survival data. *Statistical Methods in Medical Research,* 9, 161-179. doi: 10.1177/096228020000900206

Little, R.J.A., & Rubin, D.B. (2002). Statistical analysis with missing data. 2 edition John Wiley & Sons; Hoboken, NJ

Li, X., & Hedeker, D. (2012). A three-level mixed-effects location scale model with an application to ecological momentary assessment data. *Statistics in Medicine,* 31, 3192-3210. doi: 10.1002/sim.5393

Liu, L., & Huang, X. L. (2008). The use of Gaussian quadrature for estimation in frailty proportional hazards models. *Statistics in Medicine,* 27, 2665-2683. doi: 10.1002/sim.3077

Muscat, J. E., Ahn, K., Richie, J. P., & Stellman, S. D. (2011). Nicotine dependence phenotype, time to first cigarette, and risk of head and neck cancer. *Cancer,* 117, 5377-5382. doi: 10.1002/cncr.26235

Muscat, J. E., Liu, H.-P., Livelsberger, C., Richie, J. P., & Stellman, S. D. (2012). The nicotine dependence phenotype, time to first cigarette, and larynx cancer risk. *Cancer Causes & Control,* 23, 497-503. doi: 10.1007/s10552-012-9909-x

Muscat, J. E., Stellman, S. D., Caraballo, R. S., & Richie, J. P. (2009). Time to first cigarette after waking predicts cotinine levels. *Cancer Epidemiology and Prevention Biomarkers,* 18, 3415-3420. doi: 10.1158/1055-9965

Pugach, O., Hedeker, D., Richmond, M. J., Sokolovsky, A., & Mermelstein, R. (2014). Modeling mood variation and covariation among adolescent smokers: application of a bivariate location-scale mixed-effects model. *Nicotine & Tobacco Research,* 16, S151-S158. doi: 10.1093/ntr/ntt143

Reardon, S. F., Brennan, R. T., & Buka, S. L. (2002). Estimating multi-level discrete-time hazard models using cross-sectional data: neighborhood effects on the onset of adolescent cigarette use. *Multivariate Behavioral Research,* 37, 297-330. doi: 10.1207/S15327906MBR3703_1

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. Oxford: Oxford University Press.

Wang, L., & McArdle, J. J. (2008). A simulation study comparison of Bayesian estimation with conventional methods for estimating unknown change points. *Structural Equation Modeling*, *15*, 52-74. doi: 10.1080/10705510701758265

Day	N (%) asked to quit	N asked to keep smoking	0-5	6-15	16-30	31-60	≥60	not smoking
Baseline	371	480	119	115	130	165	213	109
1	302 (81.4)	394 (82.1)	70	87	104	113	130	192
2	203 (54.7)	329 (68.5)	57	77	84	72	99	143
3	158 (42.6)	295 (61.5)	45	65	67	62	89	125
4	136 (36.7)	273 (56.9)	46	59	62	55	71	116
5	130 (35.0)	239 (49.8)	40	52	51	39	72	115
6	129 (34.8)	227 (47.3)	47	53	51	36	59	110
7	129 (34.8)	227 (47.3)	43	55	50	46	49	113

Table 1. Time to first cigarette in minutes for each day of assessment

	Asked to quit N=371	Asked to keep smoking N=480	*p*
Sex			0.66
Men	133 (35.8%)	180 (37.6%)	
Women	238 (64.2%)	299 (62.4%)	
Age, mean(SD)	37.3 (10.8)	38.9 (12.3)	0.10
cpd at BL, mean(SD)	16.4 (9.7)	16.8 (8.5)	0.21
ttf at BL, median[IQR]	20 [6-60]	15 [5-60]	0.07

cpd: cigarette per day; ttf: time to first cigarette; BL: baseline

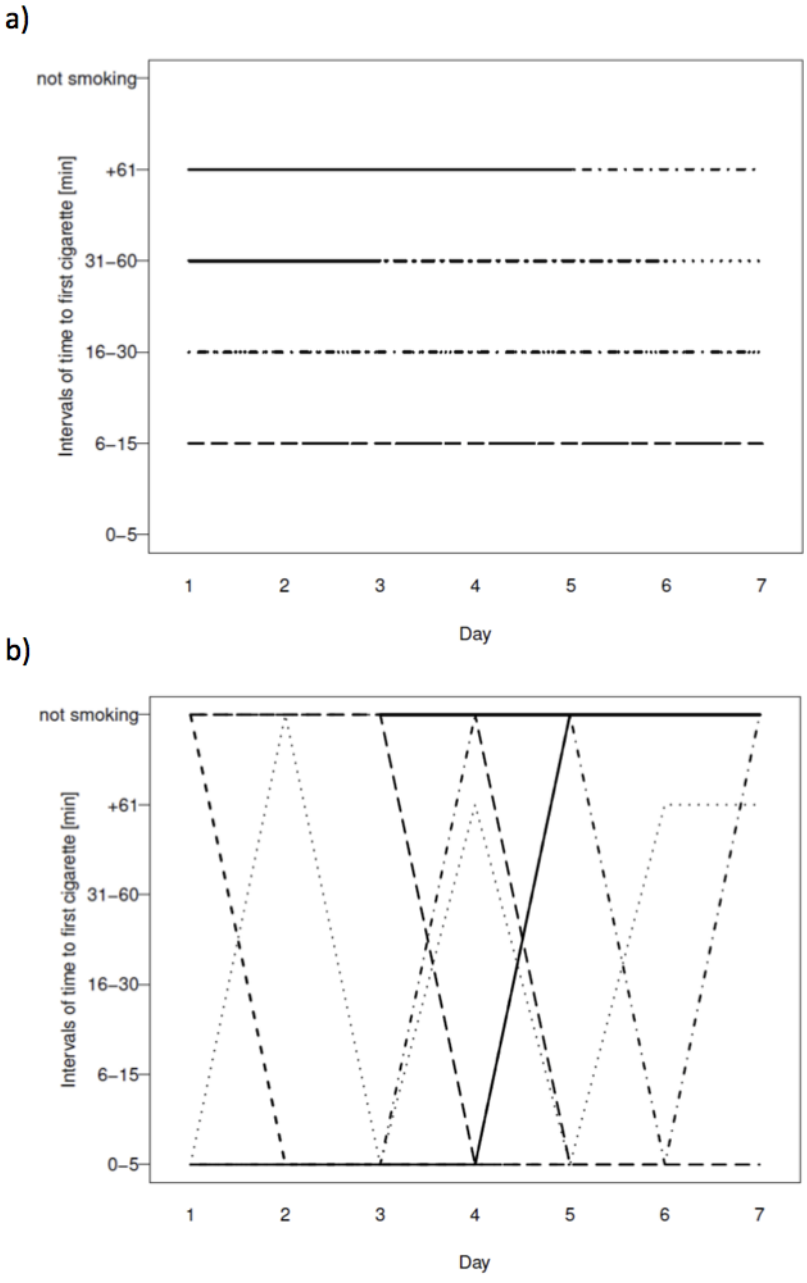
Table 2. Participant characteristics stratified by group

	Mixed-effects location			Mixed-effects location scale		
parameter	estimate	se	*p*-value	estimate	se	*p*-value
Location						
Group	4.14	0.41	<0.0001	3.88	0.46	<0.0001
Day	-0.14	0.04	0.001	-0.03	0.03	0.25
Group * Day	0.07	0.05	0.14	0.01	0.03	0.74
BS variance						
Intercept α_0	4.36	0.19	<0.0001	2.78	0.19	<0.0001
Group α_1				-0.48	0.17	0.005
WS variance						
Group τ_{group}				1.74	0.23	<0.0001
Scale						
Variance σ^2_w				3.63	0.31	<0.0001
Covariance σ^2_{vw}				-0.38	0.71	0.59
AIC	7285.8			6184.2		
BIC	7325.6			6241.6		

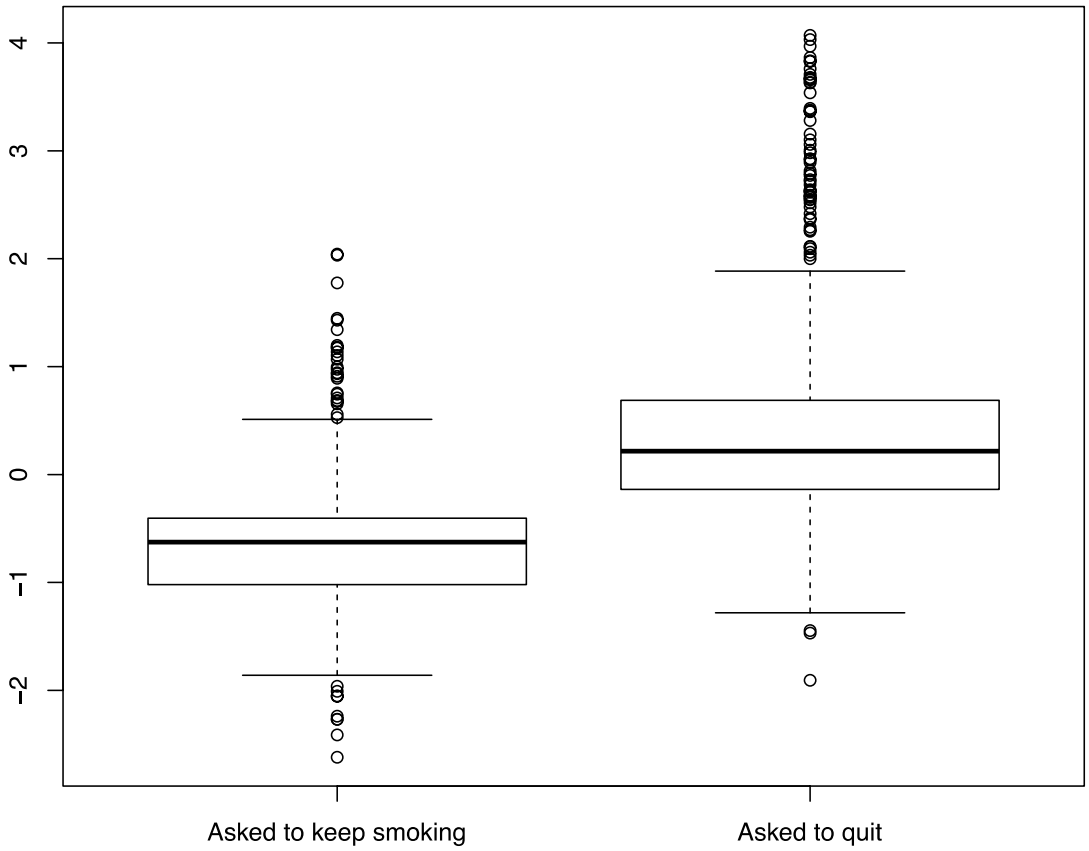
Group. One group asked to stop smoking versus one group asked to keep smoking as usual (reference group); Day. Number of day from 1 to 7. WS = within-subject variance; BS = between-subject variance; Scale = the random location scale effect; AIC = Akaike information criterion; BIC= Bayesian information criterion; α =coefficients of the covariates for BS; τ = coefficients of the covariates for WS; se: standard error.

Table 3. Proportional odds mixed effects model estimates, standard errors (se), and p-values for the random location and the random location scale models with a logit link

Figure 1. Trajectories of intervals of time to first cigarette for the 10 individuals with the lowest (Panel A) and highest (Panel B) estimate of the random scale ω_i



Supplementary figure 1: individual random scale estimates



Appendix A

Estimation

Let the vector of responses from subject i noted $\boldsymbol{Y}_i$ (of size n_i). The probability of any response pattern Y_i conditional on the random effects v_i and ω_i, is equal to the product of the probabilities of the level-1 responses:

$$\ell\left(\boldsymbol{Y}_i \mid v_i,\, \omega_i\right) = \prod_{j=1}^{ni} \prod_{t=1}^{m} Pr\left(Y_{ij} = t \mid v_i,\, \omega_i\right) \quad (11)$$

Where

$$Pr\left(Y_{ij} = t \mid v_i,\, \omega_i\right) = \psi\left(\lambda_{ijt}\right) - \psi\left(\lambda_{ij,t} - 1\right) \quad (12)$$

and ψ(.) is the logistic cumulative distribution function (cdf). The *conditional independence* assumption supposes that a subject's responses are independent conditional on the random effects, and thus allows the multiplication of each subject's responses to yield the conditional probability of the response vector. The marginal density of Y_i in the population is expressed as:

$$h\left(\boldsymbol{Y}_i\right) = \int_{v_i,\, \omega_i} \ell\left(\boldsymbol{Y}_i \mid v_i,\, \omega_i\right) f(v_i,\, \omega_i) d(v_i,\, \omega_i) \quad (13)$$

where $f\ (v_i,\, \omega_i)$ is the distribution of the random effects (in this case a bivariate normal density). While the likelihood (equation 11) represents the conditional probability, the integral of the likelihood (equation 13) indicates the unconditional probability for the response vector of subject i. Maximizing the marginal log-likelihood $\log L = \sum_i^N \log h(\boldsymbol{Y}_i)$ yields maximum marginal likelihood estimates (ML estimates). SAS PROC NLMIXED can be used to obtain the ML estimates for this model (code in appendix B).

Appendix B

Details of the options for the PROC NLMIXED analysis are provided at the end of this appendix.

```
data copd;
set copd;
if ttf>=61 then ttfCat=5;
if ttf<=60 then ttfCat=4;
if ttf<=30 then ttfCat=3;
if ttf<=15 then ttfCat=2;
if ttf<=5 then ttfCat=1;
if smoketoday eq 0 then ttfCat=6;
run;

* model with only a random intercept (ie. random location), logit link;

PROC NLMIXED data=copd GCONV=1e-12;
        PARMS gamma1=-2.9 gamma2=-2.3 gamma3=-1.8 gamma4=-1.37 gamma5=-0.02
                bGroup=1.04 bDay=-0.05 bGroupDay=0.03 sd=0.2;
         mean = bGroup*group2 + bDay*timec + bGroupDay*group2*timec + sd*u1;
clogit1 = (gamma1 + mean);
clogit2 = (gamma2 + mean);
clogit3 = (gamma3 + mean);
clogit4 = (gamma4 + mean);
clogit5 = (gamma5 + mean);
cprob1 = 1/(1+EXP(-clogit1));
cprob2 = 1/(1+EXP(-clogit2));
cprob3 = 1/(1+EXP(-clogit3));
cprob4 = 1/(1+EXP(-clogit4));
cprob5 = 1/(1+EXP(-clogit5));
        IF (ttfCat=1) THEN p = cprob1;
        ELSE IF (ttfCat=2) THEN p = cprob2-cprob1;
```

```
        ELSE IF (ttfCat=3) THEN p = cprob3-cprob2;
        ELSE IF (ttfCat=4) THEN p = cprob4-cprob3;
        ELSE IF (ttfCat=5) THEN p = cprob5-cprob4;
        ELSE IF (ttfCat=6) THEN p = 1-cprob5;
        logl = LOG(p);
        MODEL ttfCat ~ GENERAL(logl);
        RANDOM u1 ~ NORMAL(0,1) SUBJECT=idnum;
RUN;



* Mixed-effects location scale model with a logit link;

PROC NLMIXED data=copd GCONV=1e-12 tech=trureg optcheck hescal=1 qpoints=2;
        PARMS gamma1=-4.2 gamma2=-3.0 gamma3=-2.0 gamma4=-1.1 gamma5=-0.3
                bGroup=4.14 bDay=-0.14 bGroupDay=0.07 alpha0=2.0 alphaGroup=-0.7
                tauGroup=-1.3 scalevar=0.005 cov=-0.1;
            mean = bGroup*group2 + bDay*timec + bGroupDay*group2*timec+ u1;
bsvar = EXP(alpha0 + alphaGroup*group2);
wsvar = EXP(tauGroup*group2 + u2);

clogit1 = (gamma1 + mean) / SQRT(wsvar);
clogit2 = (gamma2 + mean) / SQRT(wsvar);
clogit3 = (gamma3 + mean) / SQRT(wsvar);
clogit4 = (gamma4 + mean) / SQRT(wsvar);
clogit5 = (gamma5 + mean) / SQRT(wsvar);

cprob1 = 1/(1+EXP(-clogit1));
cprob2 = 1/(1+EXP(-clogit2));
cprob3 = 1/(1+EXP(-clogit3));
cprob4 = 1/(1+EXP(-clogit4));
cprob5 = 1/(1+EXP(-clogit5));

        IF (ttfCat=1) THEN p = cprob1;
        ELSE IF (ttfCat=2) THEN p = cprob2-cprob1;
```

```
        ELSE IF (ttfCat=3) THEN p = cprob3-cprob2;
        ELSE IF (ttfCat=4) THEN p = cprob4-cprob3;
        ELSE IF (ttfCat=5) THEN p = cprob5-cprob4;
        ELSE IF (ttfCat=6) THEN p = 1-cprob5;

logl = LOG(p);
      MODEL ttfCat ~ GENERAL(logl);
      RANDOM u1 u2 ~ NORMAL([0,0],[bsvar,cov,scalevar]) SUBJECT=idnum
OUT=Copd2;
RUN;
```

Note : brief description of options used

- GCONV: specifies the relative gradient convergence criterion.
- Qpoints: number of quadrature points used during the evaluation of integrals. Setting qpoints=2 ensures that two points are used in each dimension of the random effects.
- Tech specifies the optimization technique. Tech=trureg performs a trust region optimization.
- Optcheck computes the function values of a grid of points in a ball of radius 0.1 at the starting point and 0.01 at the terminating point.
- Hescale=1 specifies the scaling version of the Hessian matrix used in TRUREG and uses the Moré scaling update.