Statistical properties of longitudinal time-activity data for use in human exposure modeling

Kristin Isaacs, Ph.D.,[1] Thomas McCurdy, M.R.P.[1] Graham Glen, Ph.D.,[2] Melissa Nysewander, Ph.D.,[2] April Errickson, M.S.L.S.[1,3] Susan Forbes, M.L.S.[3] Stephen Graham, Ph.D.,[1] Lisa McCurdy,B.Ed.[4] Luther Smith, Ph.D.,[2] Nicolle Tulve, Ph.D.,[1] and Daniel Vallero, Ph.D.[1]

[1]U.S. Environmental Protection Agency; Research Triangle Park, NC
[2]Alion Science and Technology; Research Triangle Park, NC
[3]University of North Carolina; Chapel Hill, NC
[4]Unaffiliated; Orange County, NC

Corresponding Author:
Kristin Isaacs
MD E205-02
US EPA
Research Triangle Park, NC 27711
Phone: 919-541-2785
Fax: 919-541-4787

Running Title: Properties of longitudinal time-activity data

Abstract

Understanding the longitudinal properties of the time spent in different locations and activities is important in characterizing human exposure to pollutants. The results of a four-season longitudinal time-activity diary study in eight working adults are presented, with the goal of improving the parameterization of human activity algorithms in EPA's exposure modeling efforts. Despite the longitudinal, multi-season nature of the study, participant non-compliance with the protocol over time did not play a major role in data collection. The diversity (*D*)—a ranked intraclass correlation coefficient (ICC)-- and lag-one autocorrelation (*A*) statistics of study participants are presented for time spent in outdoor, motor vehicle, residential, and other-indoor locations. Day-type (work-day versus non-work-day, and weekday versus weekend), season, temperature, and gender differences in the time spent in selected locations and activities are described, and *D* & *A* statistics are presented. The overall *D* and ICC values ranged from approximately 0.08 - 0.26, while the mean population rank *A* values ranged from approximately 0.19 - 0.36. These statistics indicate that intra-individual variability exceeds explained inter-individual variability, and low day-to-day correlations among locations. Most exposure models do not address these behavioral characteristics, and thus underestimate population exposure distributions and subsequent health risks associated with environmental exposures.

Keywords: Human activity patterns, exposure modeling, activity diaries, CHAD

Introduction

The Environmental Protection Agency (EPA) uses a sequential, time-series approach to model human exposure to several air pollutants regulated under Section 109 of the Clean Air Act, which establishes National Ambient Air Quality Standards (NAAQS). This approach requires as input realistic human time-activity (time-use) data. These data, which include information on time spent in different locations and activities by persons of different ages and genders, are used to model year-long behavior patterns for a simulated population. These behavior patterns affect both the exposures encountered by the simulated individuals and their associated ventilation rates, which together form the basis of predicted intake dose rates. To best implement the modeling approach, longitudinal time use data are needed to appropriately represent variability in an individual's locations visited and activities performed over time and space. However, there currently are few longitudinal time activity data gathered for exposure modeling purposes (1).

In the absence of comprehensive longitudinal diary data, EPA's time series exposure models must synthesize individual longitudinal time use patterns from cross-sectional data (i.e., combining diary data from similar persons), giving rise to uncertainty in the resultant information. To minimize the magnitude of uncertainty and its effect on exposure and dose estimates, EPA has developed a longitudinal diary-construction approach, called the *D* & *A* method, that reproduces realistic population properties related to exposure by selectively sampling cross-sectional diaries from EPA's Consolidated Human Activity Database (CHAD) (2). EPA models currently employing this approach include the Air Pollutants Exposure (APEX) model (3,4), the Hazardous Air Pollution Exposure Model (5), and some versions of the Stochastic Human Exposure and Dose Simulations (SHEDS) family of models (6,7).

The *D* & *A* method selectively chooses and then re-orders cross-sectional activity diary days from CHAD to 1) mimic a population-level intra- and inter-individual variance target for a user-supplied daily key diary property relevant to the exposure assessment, and 2) reproduce an appropriate day-to-day autocorrelation in a key variable (8). An example key diary variable might be the time spent by asthmatic children outside during the ozone season. Another might be the time spent in the "near-roadway" environment by an older adult with pre-existing cardiovascular problems. The *D* & *A* method is parameterized by two target properties of the population being modeled: the *D* statistic, a ratio of inter-individual to total explained variance in a key variable, and the *A* statistic, the mean day-to-day autocorrelation in the same variable. The *D* statistic can be thought of as ranked-order version of the intraclass correlation coefficient, or ICC, as explained in Glen et al. (8). Because of the relative importance of longitudinal information in its exposure modeling efforts, EPA's National Exposure Research Laboratory (NERL) is developing a database of *D* & *A* metrics for different population groups.

This paper describes results of a longitudinal activity study of eight working adults. Data were collected intermittently over the 2006-2007 time period in the Research Triangle Park, NC area. The objectives of study were to 1) identify the factors influencing time spent in different locations and activities in working adults and quantify the resulting variability; 2) quantify the intra- and inter-individual variance in time use by developing *de novo A*, *D*, and ICC metrics for this population; and 3) determine the potential for participant "protocol compliance fatigue" when an activity diary is collected for multiple periods within a year.

Methods

The Time-Activity Study

An initial sample of 9 Caucasian adults living in the Research Triangle Park, NC area was identified for this study. The sample consisted of 5 females (4 working adults, 1 non-working adult) and 4 males (all working adults) and all were co-investigators of the study. None of the participants received compensation for their efforts. We excluded the non-working female from the analyses presented here due to her unique time-use pattern associated with not working outside the home; this conforms to past practices in longitudinal time use analyses (9).

The study was staggered over a two-year period (2006-2007). Due to differences in activities and locations on different day types (e.g. work days/non-workdays or weekdays/weekend), a sampling scheme was devised that emphasized obtaining increased weekend data. The target data collection pattern was 17 consecutive days in each season: a starting Friday-Saturday pair, followed by two weeks of data collection (Sunday-Sunday), followed by a final Monday. Not all participants were available for all days in a season, but all provided at least 14 days of data for 4 of the 5 following time periods (campaigns): July 17-September 4, 2006; December 5-22, 2006; March 23-April 29, 2007; June 29-July 15, 2007; and October 25-Novermber 12, 2007. The 8 participants all provided more than the desired 28 days of data found to capture the mean time spent outdoors with a reliability coefficient $\geq$0.8 (10). They also all recorded $\geq$ 7 consecutive diary days in each season. The number of days of diary data obtained from each person is listed in Table 1. The total number of diary-days used in these analyses is 455, or approximately 57 per person. The schedule of data collection in the final sample of 8 working adults is shown in Figure 1.

A paper diary format based on Johnson (11) was used to gather daily time/activity data. Each diary was 24 h in duration, starting at midnight. Participants were instructed to fill out a new line in the diary every time he/she changed either a location or type of activity, which is known as an "event." The participants were asked to provide an entry for every event $\geq$ 1 minute in duration. Participants were also asked to record for each event whether one or more of the following circumstances occurred: 1) the subject breathed "hard" or broke a sweat; 2) window(s) or door(s) were open in an enclosed location; 3) tobacco smoking was observed in an enclosed location; 4) a combustion source was operating; and 5) solvents were being used. Participants also noted whether the day was a workday or not (regardless of its calendar day-type: weekday or weekend).

Participant Protocol Fatigue

It is well recognized in the time use research literature that the contemporaneous diary format places a significant burden on participants 12-14), but provides better detail when compared with recall methods. Also recognized is that only multiple-day diaries can properly capture activity patterns where considerable day-to-day variability exists (15). The main concern in longitudinal time use studies is compliance decreasing with increasing length of the study (15). Since our diary study had four 17-day campaigns over a 15-month time period, we were interested in determining if participant non-response occurred as the study progressed, suggesting the data gathered might be biased over time. Therefore, time-dependent trends in the number of events recorded per day were analyzed.

Time Spent in Locations and Activities

The daily time spent in a number of activities and locations were calculated for each participant. These categories were selected due to their relevance in exposure modeling. The **locations** considered included indoors (all types), inside of a residence (own and other's), other non-residential indoor locations, outdoors at home, outdoors at all other locations, and in a motor vehicle-dominated location (inside any motorized vehicle, near a roadway, in a parking lot, at a gas station, at a bus stop, in any garage, or in an auto repair shop). We refer to a person in a location of interest as a *habitué*. The **activities** considered included paid work, shopping/running errands, exercise (including walking), indoor and outdoor chores, sleeping, and cooking or preparing food. We refer to a person undertaking an activity of interest as a *doer*.

Mean daily times for each activity and location were calculated for each person, and differences due to day-type (weekday/weekend), workday status (work/non-work), gender, and temperature (>=65F versus <65F) were tested using the Wilcoxon rank sum (two-sample) test. Seasonal effects were examined with the Kruskal-Wallis test (16). In addition, descriptive statistics were obtained while treating each person-day as independent, and Kolmogorov-Smirnov (K-S) tests (16) were used to test for differences in the distribution of values for the different groups described above.

Variance and Autocorrelation Calculations

The balance of the within- and between-individual variance in the time spent in microenvironments was quantified by the intraclass correlation coefficient (ICC):

$$\text{ICC} = \sigma_{\mathbf{B}}^2 / (\sigma_{\mathbf{B}}^2 + \sigma_{\mathbf{W}}^2) \qquad (1)$$

where $\sigma_{\mathbf{B}}^2$ = inter-individual variance and $\sigma_{\mathbf{W}}^2$ = intra-individual variance (17). The denominator is total explained variance. The *D* statistic was also calculated (8). The *D* statistic is a rank-order version of the ICC, where the ranks of the time spent by each individual in microenvironments were considered instead of the raw number of minutes. Ranks were assigned to each participant for each day of the study, based on the rankings across people studied on that day. *D* was then calculated using equation 1, where the variance terms were calculated based on the assigned ranks.

D, like the ICC, is bounded by 0 and 1. As ICC and *D* approach 0, there is little inter-individual variance: intra-individual variance predominates. As ICC and *D* approach 1, within-person variance vanishes. In earlier studies, we have found the ICC statistic to be 1) quite low, generally <0.3, and 2) lower than the *D* statistic for every parameter analyzed (8). The metric (1-

ICC)/ICC gives the ratio of the intra- to inter-individual variance; it is not unusual when examining human activities for the intra-individual variance to be ≥4 times greater than the inter-individual component. The ICC values were calculated using the SAS UNIVARIATE procedure to estimate separately the between-person and within-person variance. The *D* values were calculated analogously on the scaled rank data, which were produced using a ranking algorithm in the SAS procedure IML. The estimates were adjusted for short simulation length as described in Glen et al. (8).

Since scaled ranks were used, it was a requirement that the same calendar days for each person be used to calculate *D*, as each person must have a measurement on the day in order to create the rankings. Although this is not strictly a requirement when calculating an ICC using the raw minutes, in the results reported herein, the ICC was calculated using the same calendar days as the *D* statistic. However, since some of the days in the study did not overlap for all persons, multiple estimates of *D* and ICC were calculated: more people could be included in the *D* and ICC estimates if fewer days of data were used. Giraudeau and Mary (18) provide a method for approximating the confidence intervals of the ICC as a function of the number of people and number of replicates (in this case, days) of a measurement. We used their formula to estimate which combination of N (number of people) and N_{days} (number of days) that produced the narrowest 95% confidence interval on ICC or *D*. The approximate width, *w*, of the confidence interval is given as

$$w = [2\sqrt{2}z_{(1-\alpha/2)}][1+(N_{days}-1)ICC](1-ICC)\sqrt{\frac{1}{NN_{days}(N_{days}-1)}} \qquad (2)$$

where $z_{(1-\alpha/2)}$ is the (1-α/2) percentile of the standard normal distribution. To calculate the 95% confidence interval width, w_{95}, we use α=0.05, and thus $z_{(1-\alpha/2)}$ is the 97.5th percentile of the cumulative unit normal, or 1.96.

Using equation 2, we calculated w_{95} for a range of expected ICCs for the different combinations of people and diary days that were available in this study. While using alternative combinations of days gives variable 95% CI widths over the range of expected ICCs, the differences among them were small. Therefore, we chose to use N=7 and N_{days}=26, which overall gave the narrowest confidence interval widths when the entire range of ICC or *D* was considered. Additional ICC estimates were also calculated using data on different sets of days for each person. This allowed for ICC estimates containing more days, but a *D* estimate could not be derived in this case since people could not be ranked on the same day.

The second component of the *D* & *A* method is to replicate the mean population day-to-day correlation, i.e. the lag 1 (one-day) autocorrelation (*A*), in the relevant diary property. Examples could be the day-to-day autocorrelation for time spent in residences or walking near roadways by children going to school (19). Correct characterization of the population *A* for time spent in such activities is important for correctly reproducing episodic exposures in individuals. Both Pearson (raw) and Spearman (rank) lag 1 autocorrelations in time spent in activities and locations were calculated. The values used for calculating the Spearman correlations were the ranks of days within each individual. The *A* values were calculated using the standard routines for these correlations in the SAS PROC CORR procedure.

Results

Participant fatigue: trends in the number of recorded events per day

The number of daily events, as defined above, appears as Table 1. Only one person had any day with fewer than 30 events, a compliance criterion used before in previous EPA analyses of time use data (20, 21). That person had only 2 days with <30 events out of 47 total coded days (4.3%). For the study as a whole, only 0.14% of the total sample coded <30 events on any one day. This compares favorably with CHAD as a whole, which has a 10% non-compliant rate. The median number of events recorded per day in our study is higher than all but one time use survey in CHAD--an unusual observer-coded diary of children's activities with a median of 66 events day^{-1}.

There was no clear pattern in the distribution of events recorded. Some participants' events/day followed a normal distribution while others approximated a log-normal. The trend over time in recorded events was evaluated by regressing events/day on day number. Three participants had a significantly ($p<0.05$) positive trend in their number of events recorded, and three had a negative trend; R^2 values were low even in these cases.

Descriptive statistics of time spent in locations and activities

Descriptive statistics for time spent by our sample in various locations and activities are given in Tables 2 and 3. For both tables, statistics are provided for individually-averaged data: included are the mean and standard deviation (SD) for the subsamples depicted, and their associated coefficient of variation (CV), an indicator of relative variance. The results are disaggregated by gender, day-type, workday status, temperature, and season wherever there was a significant effect of these factors.

Summary statistics for microenvironmental *locations* are given in Table 2. As these were individually-averaged data, the mean refers to the mean across people of the daily average values. There were no gender differences in time spent in various locations; however, day-type, temperature, workday status, and season were significant factors for multiple locations. Time spent in motor-vehicle dominated locations was uniform across all categories. When significant, workday status was as good as or better at discriminating time use than day-type, and much of the difference among seasons could be captured by considering temperature. Time spent outdoor in other locations was only a function of workday. When outdoor time at home and in other locations were combined in total outdoor time (not shown in table for length), there were significant differences for day-type (55 min/weekday, 120 min/weekend day), workday status (34 min/workday, 119 min/non-workday), temperature (57 min/colder day, 106 min/warmer day), and season (61 min/day in winter, 79 min/day in spring, 110 min/day in summer, 49 min/day in fall). The pattern of total outdoor and vehicle time by day-type/temperature classes for all individuals is shown in Figure 2.

Table 3 provides individually-averaged results for time spent in different *activities.* Using the Wilcoxon test, none of the descriptive categories evaluated influence time spent in exercise, personal care, or performing indoor chores. Time spent preparing food was the only variable affected by gender. Once again day-type and workday status were the most discriminating factors, affecting sleep, paid work, outdoor chores, and shopping. Paid work was the only activity that was influenced by season and temperature, with more work being performed in winter and spring, and when average temperatures were lower than 65 F.

Variance and autocorrelation estimates

Estimated values of *D* and ICC are given in Table 4. While *D* values are most useful for parameterizing EPA's longitudinal diary algorithms, ICC values are provided due to their wide use in the time use literature (10, 22). Because within-day ranking was required to calculate a *D* statistic, it was calculated using the same calendar days for each person, while ICC values were calculated using the entire set of available days. *D* values were calculated only for all days and both genders combined due to sample size issues; ICC values could be calculated for key locations and various day/gender combinations. ICC values in general had a smaller w_{95} than their corresponding *D* estimate, due to the increased number of days included in the estimate. However, the ICC includes a component of within-person variance due to weather differences or day-type effects. The ICC on workdays was always higher than on non-workdays. Both the within-person and between-person variances increased on non-workdays, but the magnitude of the increase in σ_w was much larger, resulting in a smaller ICC relative to workdays, but this difference was not significant. While females generally have lower ICC estimates, the confidence interval widths do not indicate a significant difference. Overall, both the *D* and ICC metrics are low—and their associated w_{95} confidence intervals are large, indicating significant intra-individual variability in the sample for all of the location/day-type/gender combinations evaluated.

Values of the lag-one Spearman autocorrelation, *A*, were widely spread across individuals for each variable considered. These are presented in Table 5. The value of *A* in outdoor locations ranged from 0.14-0.47. Participants with the largest *A* in outdoor time (e.g. M3, F1) also tended to have larger *A* values in other locations. Mean estimates of both the Pearson and Spearman autocorrelations across all participants are given in Table 6. This table also presents differences in *A* for workdays versus non-workdays and males versus females. The overall *A* values were similar for all considered locations/activities (ranging from 0.23 for outdoor locations to 0.29 for residences). There were significant workday category differences in the Pearson (raw) *A* for residential and other indoor locations; when the Spearman *A* was considered there was an additional difference in *A* work activities. There were no gender differences in *A* (Spearman or Pearson).

Discussion

The data we present here were designed to assess longitudinal time-activity patterns in a sample of working adults in order to assess both mean behavior and intra- and inter-person variance in time use and microenvironmental location. While ignoring intra-individual variability does not bias mean estimates of time use data, it attenuates true correlation coefficients among population parameters, and causes relational techniques—like multiple regression analyses—toward the null

(23). In one sense, intra-individual variability functions as a random variable, greatly affecting parameter distributions, particularly at their tails—the most important portion for many risk assessments (24, 25). In this section we discuss our results relative to other cross-sectional and longitudinal studies in adults, and consider the utility of the results in exposure assessment.

Participant fatigue: trends in the number of recorded events per day

Three participants had a significantly ($p<0.05$) positive trend in their number of events recorded, and three had a negative trend. However, all the R^2 values were low for all participants, explaining <10% of the total variance in events/day. A close analysis of the three "trend" participants indicated that none of them had a monotonic trend in events recorded/day for the four sampling cycles (or seasons), and that the number of events recorded per day even in the lowest cycle was well within those recorded in other contemporaneous diary studies included in CHAD (20). Given these results, we do not believe that participant fatigue would greatly influence the longitudinal time use data analysis results reported in subsequent sections. This finding is consistent with data presented in Glorieux & Minnen (26) and Schwab et al. (27), but "respondent fatigue" has been observed in other longitudinal diary studies (28).

Time spent in locations and activities

With respect to locations, at α=0.05 there were statistically significant differences in time spent in all of locations analyzed for all of distinctions evaluated (day-type, workday/non-workday, temperature class, and season of the year; see Table 1). These findings are generally consistent with the main findings of other longitudinal studies (9, 29-31), but are inconsistent with comparable analyses of cross sectional data contained in CHAD (20), pointing out potential problems of using population-weighted average data for longitudinal exposure assessments.

Time spent in selected activities shown in Table 3 indicate statistically differences in time spent in the activities for all of the distinctions evaluated (day-type, workday/non-workday, temperature class, and season; see Table 3). These findings are generally consistent with comparable activities analyzed in Wu et al. (31) and in Zuzanek & Smale (32). However, in these papers differences in how the activities were defined make direct comparisons with our findings difficult.

The only location or activity for which there was a gender difference in the individually averaged means was preparing food, and the difference was quite striking (F 31.71 ± 13.93 min versus M 4.28 ± 6.20). However, when all days were considered independently, we also found gender differences for overall distributions of residence, indoor-other, and outdoor times. This finding is roughly consistent with Echols et al. (29), Graham & McCurdy (20), Wu et al. (9,31), and Zuzanek & Smale (32). However, since there was an age discrepancy between males and females in our study, these gender differences may be confounded by age differentials.

It is clear from Figure 2 that the mean trends in day-type and temperature differences in outdoor time were maintained within individuals. The majority of the 8 participants demonstrated increased time spent outdoors on non-workdays and on warmer days. The one male participant who demonstrated more significant outdoor time on workdays habitually played sports during lunchtime and after work, demonstrating how lifestyle differences can result in uniquely different activity patterns.

Variance and autocorrelation (D & A) estimates

In general, the calculated *D* values were higher than the corresponding ICCs. This is consistent with previous observations (8) and with the hypothesis that by using ranks, the variations in everyone's behavior due to global factors such as weather are largely removed. The use of ranks versus raw data affected variables differently; for example, the *D* value for work activities was increased over the ICC more than other activities or locations, suggesting that global changes in this variable are more pronounced (which is hinted at by the larger CV values for this activity in Table 3). The mean values found in this study were similar to the ICC values presented in earlier studies, even for quite dissimilar populations. These studies include 1) the Frazier et al. (22) analysis of older adult data from two US locations; 2) the Glen et al. (8) reanalysis of data from school-age children in the Harvard Southern California Chronic Ozone Exposure (HSCOS) study (33); 3) the Wu et al., (31) analysis of parent's time use data; and 4) the Xue et al. (10) analysis of the HSCOS data. All of these longitudinal time use studies, and analyses of physical activity data, demonstrate significant within-person variation in time spent in major (and aggregated) locations, such as time spent in transit, outdoors, shopping areas, and—surprisingly--even in home and work locations. The ICC values from these studies are in the range of 0.15-0.40, even though disparate age/gender cohorts were evaluated. This is an important finding, as discussed below.

Despite the general agreement of mean ICC and D values with previous findings the w_{95} confidence intervals were quite large due to the small sample size. This imposes obvious limitations on the application of these results directly to EPA efforts. However, it is clear in most cases that at least some measureable amount of within-individual variability is present (even in this homogeneous population), and in come cases using the mean values presented here may preferable to ignoring such variance altogether. We admit that larger studies will be required to confirm any recommended "target" values of *D* or ICC for time spent in activities/locations. We anticipate in the near future analyzing additional longitudinal datasets from available exposure studies with larger N (9,34,35) that are currently being added to CHAD.

The mean population *A* (autocorrelation rank) values in this study shown in Table 6 were also similar to those found for the children in the Harvard Southern California study (8) for outdoor time (0.24 in this study versus 0.22), indoor times (0.20 for residences and 0.23 for other indoor versus 0.22 for all indoor). They also are similar to those seen in the Frazier et al. (2009) analysis of time use by older adults, except that group has a much higher *A* for the residential location (≥0.50). Since *A*'s that high have never been seen in any other analysis that we know of, that value must be a direct result of older adults spending so much time at home and not travelling much.

EPA's *D* & A method has the potential to allow targeting of *A* for different day-types or by gender. Values for *A* for workdays versus non-workdays and males versus females are reported in Table 6. There were several significant differences between day-types, both for the raw *A* values and those based on the daily rank of the variable. Values of *A* for males and females were similar. Thus the current study does not indicate a need for a more complex implementation of the autocorrelation algorithm that considers gender. The longitudinal data presented here aid in understanding the importance of intra-individual variability over time. Even within day-types, the ratio of between-person variance to total variance was quite low (Table 5). Even considering the larger $w_{95,}$ significant within-person variance in time spent in different locations is likely, even when the influence of other factors has been removed.

It should be noted that this is a small study of a relatively homogenous population of working professional adults. While the values found here are similar to those found earlier in a variety of population subgroups, further characterization of these properties in other population cohorts and in other regions of the U.S. would be desirable. Since employment status, weather conditions, age and gender, and socioeconomic factors impact longitudinal patterns and their properties, large sample size studies of diverse population cohorts are needed to fully explicate the ICC, *D*, and *A* statistical targets used in EPA's time series exposure models.

Using ICC, D and A: Exposure modeling and assessment

The ICC, D and A values estimated here can aid in understanding and quantifying the variability in longitudinal time use in exposure assessment. Longitudinal assessment of exposure typically involves constructing some type of time-series of time spent in microenvironments, either for individuals or cohorts. These results demonstrate the importance of targeting both mean behavior and intrapersonal variability. While EPA's current longitudinal diary algorithm (8) is parameterized with *D*, the raw ICCs reported here can also provide guidance. Attempts at modeling the time spent in microenvironments as a function of day-type, season, and other demographic or temporal variables have been undertaken, but the predictive strength of the resulting models have typically been low (10), because too many factors--many perhaps related to lifestyle or occupation--have not yet been quantified. Therefore, many approaches used by EPA and other organizations (e.g., 36) have focused on sampling real activity and location data from CHAD or other data sources. These methods include sampling a new diary for every day in the simulation, or sampling one diary per year for each individual for each season/day-type combination, or some combination of these approaches. The reported ICC values can be used to assess different sampling strategies. The fact that the ICCs for specific day-types and temperatures are relatively low indicate that a single sample from a cohort-specific diary pool may not be adequate for quantifying variability in temporal patterns. Simulated ICC values will be influenced by number of different diaries selected for each person, and how often they are repeated over the simulation period. Accurately partitioning within-person and between-person variation in time spent in the microenvironment can avoid over- or under-estimation of individual variability in exposures.

In assessing longitudinal data for use in exposure modeling, the ability to characterize the mean behavior of the population is critical. The number of days required to estimate mean behavior is a function of the observed ICC, and can be predicted by the Spearman-Brown prophecy formula:

$$N_{days} = \frac{[R(1-\mathrm{ICC})]}{[\mathrm{ICC}(1-R)]} \tag{3}$$

where *R* is a target reliability in the estimate of mean behavior (for example, time spent in microenvironments). EPA has previously shown that $\geq$28 days of time use data and 7+ consecutive days per season are needed to properly characterize children's time spent outdoors using R=0.8 (10). Equation 3 can be used to demonstrate an advantage of parameterizing longitudinal algorithms using *D* rather than ICC. In general, rankings of time should be a more stable variable than raw time, since the variation of everyone's behavior due to the weather and other factors (like holidays) is largely removed. Thus, *D* will likely be larger than the raw ICC because the within-person variance has been decreased. This was generally the case in this study (Table 4) and previous ones (8). Using *D*=0.26 and ICC=0.07 (the values for outdoor time in this study), equation 3 predicts that only 12 days of data are needed to achieve the same reliability (R= 0.8) for the rank of outdoor time that one gets with 36 days of data for raw outdoor time. Thus, the mean rank of behavior for an individual can be assessed using fewer days of data, and thus more studies may be available for charactering such metrics.

The confidence in the measured value of ICC itself is dependent upon the number of days and the number of people being studied (as predicted by equation 2). An example is shown in Figure 3, which shows (for different values of ICC or *D*) the N and N_{days} required to achieve a w_{95} of 0.05. As N increases, the confidence in ICC increases, and thus a smaller N_{days} is required to achieve a desirable w_{95}. Figure 3 illustrates that ICC and *D* are collective properties that may be estimated fairly well even when the mean properties for individuals cannot be reliably measured. For example, if 400 people are studied, a good w_{95} can be achieved in 10 days for an ICC (or *D*) of 0.1, even though equation 2 indicates that 36 days are needed to estimate mean time spent in microenvironments. This is encouraging since longitudinal studies are rare, and keeping accurate activity diaries results in significant participant burden, oftentimes resulting in fewer days of data collection than are required for estimating mean behavior (for example, 37, 38).

Together, equations 2 and 3 can be used to evaluate existing data sources or design future studies for characterizing ICC and *D*. Datasets having large numbers of people and few days may be quite useful for getting estimates of these parameters, even if not enough days are available to reliably estimate mean time spent in locations or activities. The opposite is also true: some studies may be adequate for estimating time spent in microenvironments and activities even though the confidence intervals in w_{95} are quite wide. How well ICC or *D* and *A* need to be characterized in EPA's models due to sensitivity concerns is an ongoing area of research.

Conclusions

This paper presents a statistical analysis of time spent in various locations and activities in a sample of working adults in Research Triangle Park, NC. Approximately 57 days of activity data were collected from each person over all four seasons. The objectives of study were to 1) determine the potential for participant "protocol compliance fatigue" in keeping activity diaries, 2) identify the factors influencing time spent in different locations and activities in working adults and quantify the resulting variability and lastly 3) quantify metrics describing intra- and inter-individual variance and autocorrelation in activity for use in exposure modeling and assessment.

Based on the results presented herein, we found little evidence of participant fatigue in this study. This is encouraging, as the paper diary collection methods were somewhat burdensome. It is hoped that new data collection technologies (e.g. using smartphones) will further reduce the burden of diary-keeping and allow for even longer collection periods in individuals.

In this study of working adults, the only gender difference in time spent in activities and microenvironments was time spent preparing food. Overall, seasonal effects could be accounted for by considering temperature differences, and considering workday/non-workday as opposed to weekday/weekend differences improved characterization of behavior.

The intraclass correlation coefficients for both raw times (ICC) and ranks of times among individuals (*D*) were assessed. The ICC and *D* values were typical of those seen in other studies for times spent in microenvironments and activities. The values were typically on the order of 0.15-0.40, indicating a high degree of within-person variability even for this fairly homogeneous population of professional adults. These values can be used to help parameterize EPA's or other similar longitudinal exposure models; methods for assessing data from future longitudinal activity studies were presented. Both the number of people studied and the number of days of data collected for each individual are crucial when determining the utility of a particular longitudinal study in estimating *D* & *A*, as the width of the 95^{th} confidence interval on the ICC (or *D*) is a function of both of these quantities. The equations presented herein can be used to assess other available longitudinal human activity datasets, and serve as a tool in the optimal design of future longitudinal time-use studies.

ICC and *D* values on the order found in this study and the others mentioned above—0.15 to 0.40 overall—indicate that intra-individual variance is between 2 and 5 times as large as inter-individual variability for fairly homogeneous population subsets. Intra-individual variability is largely ignored in the exposure modeling community, particularly in those models that focus on time-averaged exposure and dose metrics. In a sense, this essentially is analogous to ignoring uncertainty in an important aspect of exposure modeling: human time/activity patterns. In addition, many exposure models are also deterministic in nature, using point estimates of the time spent in various locations, and not varying those estimates by day-type, season of the year, outdoor temperature regime, etc. Doing so basically ignores inter-individual variability except in a crude sense when the modeled population is disaggregated into gender and broad age groupings. If intra- and inter-individual variability in time use (and other important inputs) is ignored or under-defined, the resulting exposure and dose distributions will be much narrower than warranted by the data, particularly at the "high-end" of the distributions where health risks are most important (39, 40). Therefore, the protection received from control strategies that are used as input scenarios to regulatory exposure/dose modeling efforts may provide misleading information. We hope that the information presented here adds to the knowledge base concerning intra- and inter-individual variability in how people spend their time and in what locations.

Disclaimer

The United States Environmental Protection Agency (EPA) through its Office of Research and Development conducted the research described in this paper. It has been subjected to Agency review and approved for publication.

References

1. Hertz-Picciotto I, Cassady D, Lee K, Bennett DH, Ritz B, Vogt R. Study of Use of Products and Exposure-Related Behaviors (SUPERB): study design, methods, and demographic characteristics of cohorts. Environ Health 2010; 9: 54.
2. McCurdy T, Glen G, Smith L, Lakkadi Y. The National Exposure Research Laboratory's consolidated human activity database. J Expos Anal Environ Epidemiol 2000; 10: 566-578.
3. US EPA. Total Risk Integrated Methodology (TRIM) Air Pollutants Exposure Model Documentation (TRIM.Expo/APEX, Version 4.3). Volume 1: Users Guide (EPA-452/B-08-001a). Research Triangle Park, NC: U.S. Environmental Protection Agency; 2008a.
4. US EPA. Total Risk Integrated Methodology (TRIM) Air Pollutants Exposure Model Documentation (TRIM.Expo/APEX, Version 4.3). Volume 2: Technical Support Document (EPA-452/B-08-001b). Research Triangle Park, NC: U.S. Environmental Protection Agency; 2008b.
5. Rosenbaum A, Huang M. The HAPEM6 User’s Guide. Research Triangle Park, NC: U.S. Environmental Protection Agency; 2007.
6. Zartarian V, Xue J, Glen G, Smith L, Tulve N, Tornero-Velez R. Quantifying children's aggregate (dietary and residential) exposure and dose to permethrin: application and evaluation of EPA's probabilistic SHEDS-Multimedia model. J Expos Sci Environ Epidemiol 2012; 22: 267–273.
7. Xue J, Zartarian V, Liu S, Geller A. Methyl Mercury Exposure from Fish Consumption in Vulnerable Racial/Ethnic Populations. Sci Tot Env 2012; 414: 373–379.
8. Glen G, Smith L, Isaacs K, McCurdy T, Langstaff, J. A new method of longitudinal diary assembly for human exposure modeling. J Expos Sci Environ Epidemiol 2008; 18: 299-311.

9. Wu X, Fan Z, Ohman-Strickland P. Time-location patterns of a population living in an air pollution hotspot. J Environ Public Health [Internet] 2010; Available from: http://www.hindawi.com/journals/jeph/2010/625461/.
10. Xue J, McCurdy T, Spengler J, Ozkaynak H. Understanding variability in time spent in selected locations for 7-12-year old children. J Expos Anal Environ Epidemiol 2004; 14: 222-233.
11. Johnson T. Human Activity Patterns in Cincinnati, Ohio. Palo Alto, CA: Electric Power Research Institute; 1989.
12. Ås D. Studies of time use: problems and prospects. Acta Sociol 1978; 21: 125-134.
13. Kan MY, Pudney S. Measurement error in stylised and diary data on time use. Sociol Meth 2008; 38: 101-132.
14. Robinson JP. How Americans Use Time. New York: Praeger Publishers; 1977.
15. Harvey AS. Guidelines for time use data collection. Social Indic Res 1993; 30: 197-228.
16. Conover WJ. Practical Nonparametric Statistics. 2nd ed. New York: John Wiley & Sons; 1980.
17. Shrout PE, Fleiss JL. Intraclass Correlations: Uses in Assessing Rater Reliability. Psych Bull 1979; 2: 420-428.
18. Giraudeau B, Mary JY. Planning a reproducibility study: how many subjects and how many replicates per subject for an expected width of the 95 per cent confidence interval of the intraclass correlation coefficient. Stat Med 2001; 20: 3205-3214.
19. Xue J, McCurdy T, Burke J, Bhaduri B, Liu C, Nutaro J et al. Analyses of school commuting data for exposure modeling purposes. J Expos Sci Environ Epidemiol 2010: 20: 69-78.

20. Graham S.E., and McCurdy T. Developing meaningful cohorts for human exposure models. J Expos Anal Environ Epidemiol 2004; 14: 23-43.

21. McCurdy T, Graham SE. Using human activity data in exposure models: analysis of discriminating factors. J Expos Anal Environ Epidemiol 2003; 13: 294-317.

22. Frazier EL, McCurdy T, Williams R, Linn WS, George BJ. Intra- and inter-individual variability in location data for two U.S. health-compromised elderly cohorts. J Expos Sci Environ Epidemiol 2009; 19: 580-592.

23. Ridley K, Olds T, Hands B, Larkin D, Parker H. Intra-individual variation in children's physical activity patterns: implications for measurement. J Sci Med Sport 2009; 12: 568-572.

24. Hattis DL. Distributional analyses for children's inhalation risk assessments. J Tox Environ Health 2008; 71: 218-226.

25. Hattis DL, Anderson EL. What should be the implications of uncertainty, variability, and inherent 'biases'/'conservatism' for risk management decision-making. Risk Anal 1999; 19: 95-107.

26. Glorieux I, Minnen J. How many days? A comparison of the quality of time-use data from 2-day and 7-day diaries. Inter J Time Use Res 2009; 6: 314-327.

27. Schwab M, McDermott A, Spengler JD. Using longitudinal data to understand children's activity patterns in an exposure context: data from the Kanawha County health study. Environ Inter 1992; 18: 173-189.

28. Stopher P, Greaves S, Clifford E. Multi-day household travel surveys: sampling issues. 30th Australasian Transport Research Forum 2007; 1-12.

29. Echols SL, MacIntosh DL, Hammerstrom KA, Ryan PB. Temporal variability of microenvironmental time budgets in Maryland. J Expos Anal Environ Epidemiol 1999; 9: 502-512.

30. Rodes C, Lawless P, Thornburg J, Croghan C, Vette A, Williams R. DEARS particulate matter relationships for personal, indoor, outdoor, and central site settings for a general population. Atmos Environ 2010; 4: 1386-1399.

31. Wu X, Bennett DH, Lee K, Cassady DL, Ritz B, Hertz-Picciotto I. Longitudinal variability of time-location/activity patterns of population at different ages: a longitudinal study in California. Environ Health [Internet] 2011; 10: 80. Available from: http://www.ehjournal.net/content/10/1/80.

32. Zuzanek J, Smale BJ. Life-cycle variations in across-the-week allocation of time to selected daily activities. Loisir et Société 1993; 15: 559-586.

33. Geyh AS, Xue J, Ozkaynak H, Spengler JD. The Harvard Southern California Chronic Ozone Exposure Study: assessing ozone exposure of grade-school-age children in two Southern California communities. Environ Health Perspect 2000; 108(3): 265-270.

34. Phillips MJ, Rodes CE, Thornburg JW, Whitmore R, Vette A, Williams R. Recruitment and Retention Strategies for Environmental Exposure Studies: Lessons from the Detroit Exposure and Aerosol Research Study. RTI Press publication No. MR-0021-1011. Research Triangle Park, NC: RTI Press; 2010. Available at: http://www.rti.org/pubs/mr-0021-1011-phillips.pdf.

35. Susan Lyon Stone SL, Graham S, Pekar Z, Mansfield C, Depro B, Isaacs K et al. Using EPA's Air Quality Index (AQI) to reduce cardiovascular morbidity and mortality in older adults. American Public Health Association 140th Annual Meeting, San Francisco. Abstract 260301; 2012.

36. Loh MM, Houseman EA, Levy JI, Spengler JD, Bennett DH. Contribution to volatile organic compound exposures from time spent in stores and restaurants and bars. J Expos Sci Environ Epidemiol 2009; 19: 660-673.

37. Freeman NC, Lioy PJ, Pellizzari E, Zelon H, Thomas K, Clayton A et al. Responses to the region 5 NHEXAS time/activity diary. National Human Exposure Assessment Survey. J Expos Anal Environ Epidemiol 1999; 9: 414-426.

38. Williams R, Suggs J, Rea A, Leovic K, Vette A, Croghan C, et al. The Research Triangle Park particulate matter panel study: PM mass concentration relationships. Atmos Environ 2003; 37: 5349-5363.

39. Hattis D, Russ A, Goble R, Banati P, Chu M. Human interindividual variability in susceptibility to airborne particles. Risk Anal 2001; 21: 585-599.

40. McCurdy T. Modeling the dose profile in human exposure assessments: ozone as an example. Rev Tox: In Vivo Tox Risk Assess 1997; 1: 3-23.

Figure Legends

Figure 1. Pattern of seasonal data collection in 8 working adults, 2006-2007.

Figure 2. Time spent outdoors and in vehicles by all participants for work days and non-work days in two temperature categories (T<65 F and T≥65 F).

Figure 3. Number of People and Diary Days Required to Achieve a w_{95} of 0.05 for Different Values of ICC (or D).

Table 1. Analysis of the number of diary events per recorded day

Participant (M=Males, F=Females)	Age (y)	Days of Data	Number of Events Recorded Per Day by the Participants				Significant Trend? (p for Days<0.050)			
			Mean	SD	CV (%)	Range	R^2 (%)	b	p	
F1	35	56	52	7	14.4	39 - 66	3.84	0.100	0.079	
F2	36	53	56	9	16.8	41 - 82	0.02	0.010	0.918	
F3	37	61	42	6	14.0	30 - 53	9.18	0.110	0.010	*
F4	39	58	54	7	12.4	38 - 74	2.66	0.083	0.115	
M1	52	54	41	5	13.0	31 - 57	8.00	-0.105	0.022	*
M2	54	52	48	6	13.7	37 - 70	0.17	-0.018	0.771	
M3	54	47	36	4	12.2	27 - 46	4.18	0.155	0.087	
M4	66	74	52	7	12.9	40 - 68	7.72	-0.023	0.044	*

Notes & symbols:

CV=Coefficient of variation (SD/mean)
SD=Standard deviation
R^2=Coefficient of determination
b=Slope of the regression line
*=Statistically significant from 0.0 at $p<0.05$

Table 2. **Selected descriptive statistics for individually-averaged data: time spent in selected locations in minutes/day (n=8 people)**

Location	Category	Minutes/Day Mean (SD)	CV (%)	Test p-value
Indoor - Residence				
Day-type:	Weekday	811.3 (74.8)	9.2	0.001
	Weekend	1049.3 (132.3)	12.6	.
Workday:	Workday	787.2 (63.7)	8.1	0.002
	Non-Workday	1007.8 (141.5)	14.0	.
Temperature:	Average Temp <65 F	864.5 (106.7)	12.4	0.046
	Average Temp >=65 F	927.3 (117.9)	12.7	.
Indoor – Other				
Day-type:	Weekday	487.4 (51.3)	10.5	< 0.001
	Weekend	169.8 (65.8)	38.7	.
Workday:	Workday	543.4 (49.3)	9.1	< 0.001
	Non-Workday	207.3 (71.5)	34.5	.
Temperature:	Average Temp <65 F	414.0 (48.7)	11.8	0.005
	Average Temp >=65 F	324.2 (72.7)	22.4	.
Season:	Winter	423.5 (34.1)	8.0	0.012
	Spring	365.9 (68.3)	18.7	.
	Summer	343.3 (80.3)	23.4	.
	Fall	288.1 (122.9)	42.7	.
Outdoor – Residence				
Day-type:	Weekday	20.6 (13.0)	63.1	0.036
	Weekend	63.4 (50.4)	79.5	.
Workday:	Workday	10.8 (6.6)	61.6	0.009
	Non-Workday	59.6 (38.6)	64.7	.
Temperature:	Average Temp <65 F	16.3 (16.5)	101.1	0.036
	Average Temp >=65 F	56.9 (35.3)	62.1	.
Season:	Winter	11.3 (10.0)	88.3	0.017
	Spring	47.1 (33.8)	71.8	.
	Summer	54.3 (41.9)	77.3	.
	Fall	29.5 (25.9)	88.0	.
Outdoor - Other				
Workday:	Workday	26.6 (24.6)	92.5	0.006
	Non-Workday	65.0 (27.2)	41.8	.
Motor-Vehicle Dominated		85.8 (29.6)	34.6	.

Symbols:

CV=Coefficient of variation (SD/Mean)

n=Number of people who contributed data

Test=Wilcoxon test for 2-way comparisons; Kruskal-Wallis test for season

SD=Standard deviation

Table 3. Selected descriptive statistics for individually-averaged data: time spent in different activities in minutes/day.

Activity	Category	n (People)	Mean (SD)	CV (%)	Wilcoxson Test[a] p-Value
Sleep					
Day-type:	Weekday	8	458.1 (35.0)	7.6	0.027
	Weekend	8	506.3 (38.2)	7.5	.
Workday:	Workday	8	448.4 (35.4)	7.9	0.006
	Non-Workday	8	503.5 (33.8)	6.7	.
Work					
Day-type:	Weekday	8	374.9 (51.8)	13.8	< 0.001
	Weekend	8	9.3 (11.1)	118.7	.
Workday:	Workday	8	469.0 (16.0)	3.4	< 0.001
	Non-Workday	8	26.7 (18.6)	69.8	.
Temperature:	Average Temp <65 F	8	268.8 (70.4)	26.2	0.012
	Average Temp >=65 F	8	204.4 (27.7)	13.6	.
Season:	Winter	8	282.6 (105.4)	37.3	0.029*
	Spring	8	225.9 (41.6)	18.4	.
	Summer	8	214.0 (41.0)	19.2	.
	Fall	6	185.0 (105.7)	57.2	.
Personal Care		8	56.8 (20.1)	35.3	.
Shop/Run Errands					
Workday:	Workday	8	6.4 (3.7)	57.8	0.036
	Non-Workday	8	15.2 (12.0)	79.1	.
Prepare Food					
	Female	4	31.7 (13.9)	43.9	0.021
	Male	4	4.3 (6.2)	145.0	.
Indoor Chores		8	66.1 (37.6)	56.9	.
Outdoor Chores					
Workday:	Workday	8	4.7 (5.3)	113.0	0.009
	Non-Workday	8	30.1 (36.7)	121.9	.

Symbols & Notes:

CV=Coefficient of variation (SD/Mean)
n=Number of people who contributed data
SD=Standard deviation
[a]Test that the *ranks* of two ordered samples are identical with an $\alpha=0.05$.
*p-value for season from the Kruskal-Wallis ranked-order analysis of variance (at $\alpha=0.05$)

Table 4. Estimates of D and ICC for several key locations/activities.

Key Locations	*D*[1]		ICC[2]									
	All Days[3]		All Days		Non-Workdays[3]		Workdays[3]		Females[4]		Males[5]	
	mean	w_{95}	mean	W_{95}	mean	w_{95}	mean	w_{95}	mean	w_{95}	mean	w_{95}
Outdoors (Total)	0.26	0.45	0.07	0.17	0.10	0.24	0.32	0.48	0.04	0.16	0.17	0.43
Indoor - Residence	0.20	0.40	0.12	0.25	0.20	0.37	0.37	0.51	0.01	0.08	0.22	0.52
Indoor – Other	0.08	0.23	0.02	0.08	0.08	0.21	0.24	0.42	0.00	0.01	0.04	0.16
Motor Vehicle Dominated	0.13	0.3	0.08	0.18	0.08	0.21	0.29	0.47	0.01	0.08	0.26	0.57
Work	0.22	0.42	0.00	0.04	0.04	0.15	0.08	0.23	0.02	0.11	0.00	0.01

Symbols & Notes:

D=Diversity statistic; a rank-ordered ICC

ICC=Intraclass correlation coefficient

n=Number of subjects

N_{days}=Number of days of data per subject

w_{95}=Width of the 95th percentile confidence interval about the mean

1. n=7; N_{days}=26, derived using the same calendar days for each person
2. Calculated using all available days of data
3. Workdays: n = 8 and N_{days} = 20, Non-workdays: n = 9 and N_{days} = 26
4. n = 4 and N_{days} = 53
5. n = 4 and N_{days} = 48

Table 5. Day-to-day (lag-one) Spearman (rank) autocorrelation values for selected locations & activities

Participant	Outdoor Locations	Residential Locations	Other-Indoor Locations	Motor-vehicle Locations	Work Activities
F1	0.34	0.50	0.41	0.20	0.34
F2	0.32	0.33	0.18	0.11	0.08
F3	0.22	0.11	0.18	0.15	0.27
F4	0.31	0.54	0.57	0.42	0.46
M1	0.17	0.39	0.33	0.12	0.30
M2	0.12	-0.20	-0.10	0.02	0.04
M3	0.47	0.28	0.25	0.33	0.24
M4	0.14	-0.04	0.07	0.11	0.36

Table 6. Estimates of the population mean lag-one autocorrelation (*A*).

Key Variable	Mean *A* (Raw; Pearson)					Mean *A* (Rank; Spearman)				
	All	**Non-Workdays**	**Workdays**	**Females**	**Males**	**All**	**Non-Workdays**	**Workdays**	**Females**	**Males**
Outdoor Locations	0.23	0.25	0.23	0.24	0.23	0.24	0.25	0.25	0.20	0.27
Residential Locations	0.29	0.40	0.10	0.26	0.31	0.2	0.29	0.10	0.17	0.21
Other-Indoor Locations	0.27	0.36	0.17	0.25	0.28	0.23	0.27	0.13	0.21	0.25
Motor Vehicle Locations	0.22	0.21	0.00	0.17	0.25	0.19	0.24	0.01	0.15	0.20
Work Activities	0.26	0.11	0.00	0.23	0.28	0.36	0.46	0.00	0.48	0.27

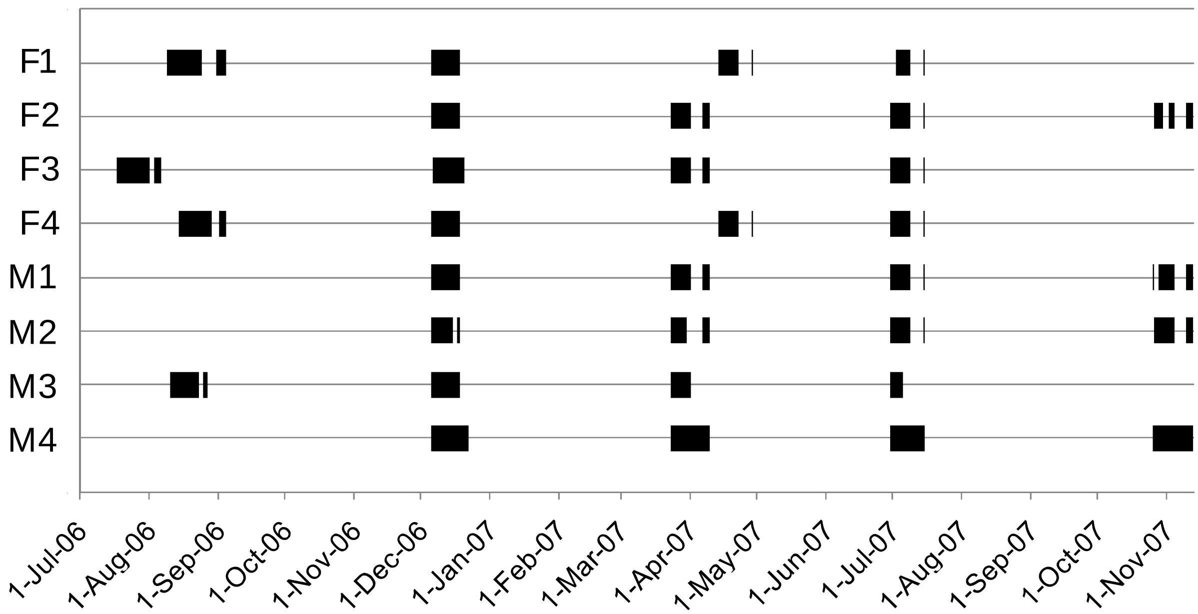

F1
F2
F3
F4
M1
M2
M3
M4
1-Jul-06
1-Aug-06
1-Sep-06
1-Oct-06
1-Nov-06
1-Dec-06
1-Jan-07
1-Feb-07
1-Mar-07
1-Apr-07
1-May-07
1-Jun-07
1-Jul-07
1-Aug-07
1-Sep-07
1-Oct-07
1-Nov-07

Figure 1.

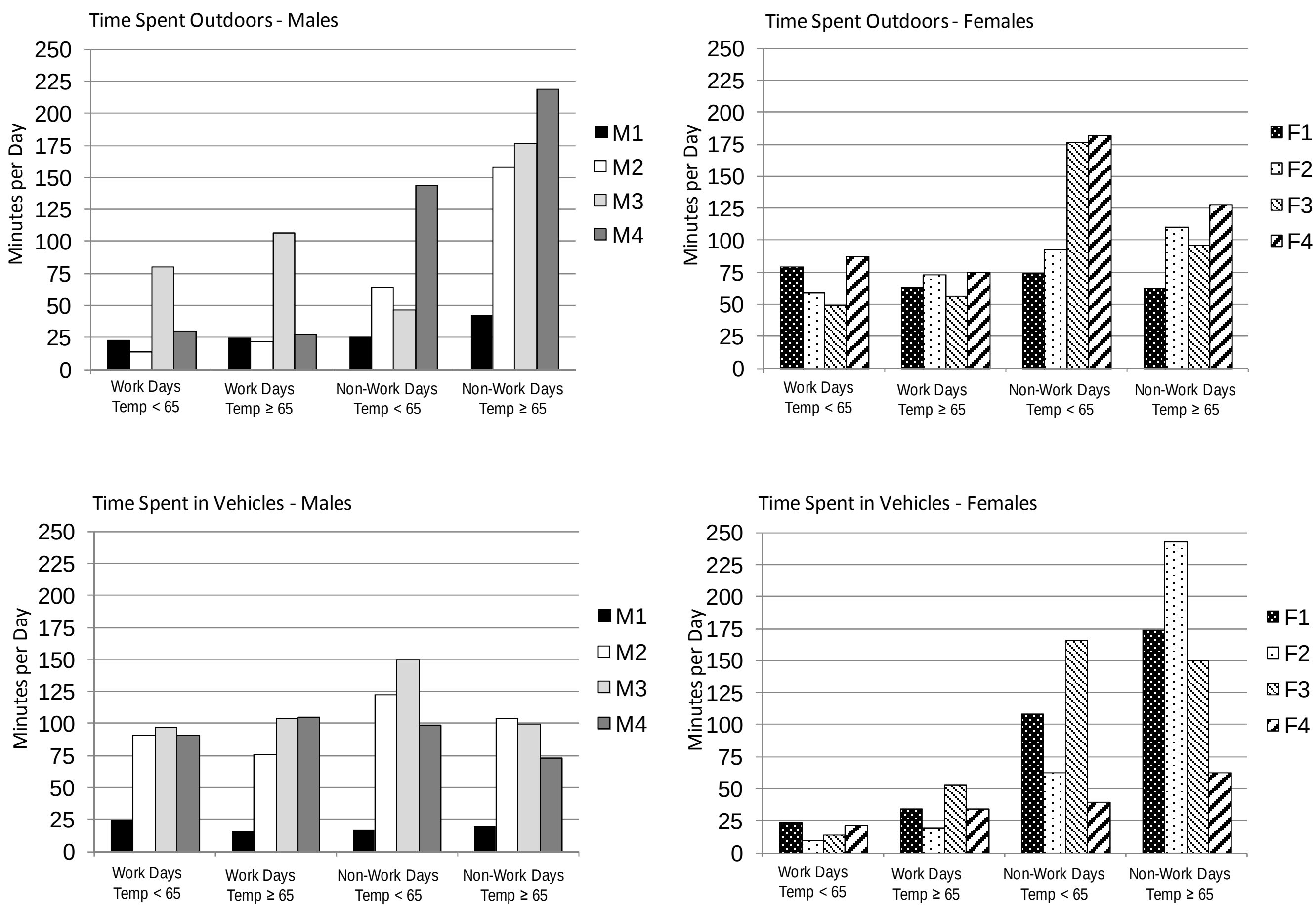
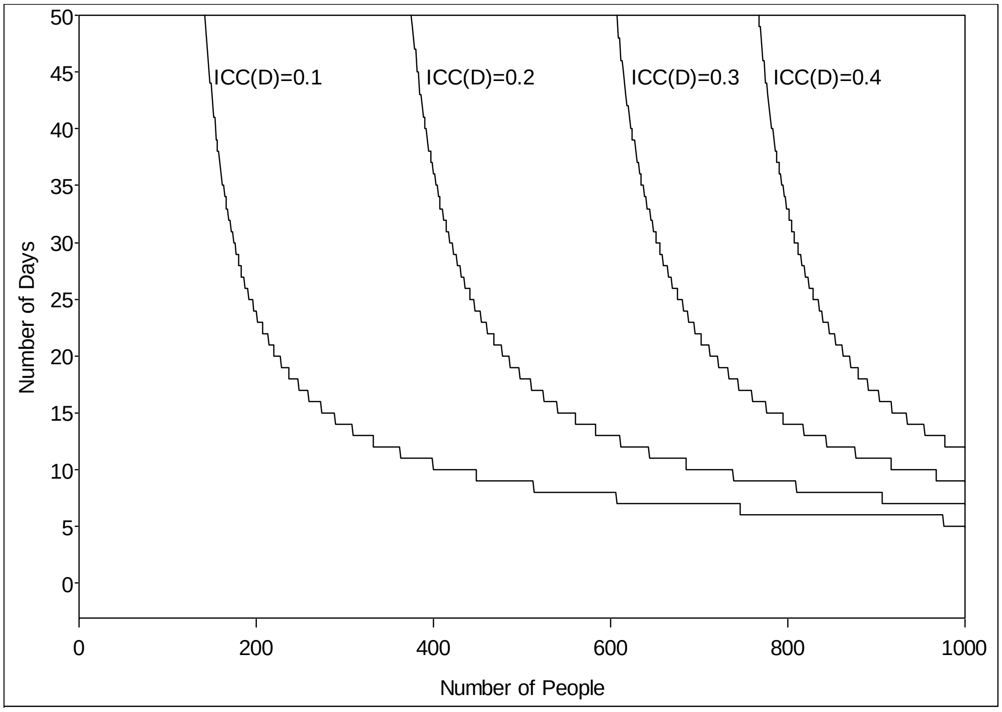

Time Spent Outdoors - Males
Time Spent Outdoors - Females
Time Spent in Vehicles - Males
Time Spent in Vehicles - Females
Minutes per Day
M1
M2
M3
M4
F1
F2
F3
F4
Work Days Temp < 65
Work Days Temp ≥ 65
Non-Work Days Temp < 65
Non-Work Days Temp ≥ 65

Figure 2.



50
45
40
35
30
25
20
15
10
5
0
Number of Days
ICC(D)=0.1
ICC(D)=0.2
ICC(D)=0.3
ICC(D)=0.4
0
200
400
600
800
1000
Number of People

Figure 3.