Author's response to reviews

Title: Longitudinal Variability of Time-Location/Activity Patterns of Population at Different Ages

Authors:

Xiangmei Wu (maywu@ucdavis.edu)
Deborah H Bennett (dhbennett@ucdavis.edu)
Kiyoung Lee (cleanair@snu.ac.kr)
Diana L Cassady (dlcassady@phs.ucdavis.edu)
Beate Ritz (britz@ucla.edu)
Irva Hertz-Picciotto (ihp@ucdavis.edu)

Version: 3 Date: 23 August 2011

Author's response to reviews: see over
Reviewer 1

Title: Longitudinal Variability of Time-Location/Activity Patterns of Population at Different Ages

Version: 1 Date: 31 March 2011

Reviewer: Thomas McCurdy

Reviewer's report: See the attached file.

Level of interest: An article of importance in its field

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I am a federal employee whose salary is paid for by the tax-payers. I have no competing interests, financial or otherwise. Therefore, I declare that I have no competing interests.

Comments on Wu et al. paper

1. The data obtained in this study are valuable and add considerably to the literature on longitudinal time use in a heterogeneous population. The paper is a valuable addition to the body of exposure-oriented human activity data regardless of my comments that follow. The authors tried a number of ways to obtain the data, which is a plus. The data seem to be consistent with previous studies, and appear to be analyzed in an innovative and meaningful way. In my opinion, the paper is worthy of publication--with "minor essential revisions" relating to the exposition (sentence structure and grammar).

2. The distinction made between occurrences (O) and duration (D) is a good one that is followed through in the analyses of data. This distinction allows a researcher to use a conditional probability approach to using time/activity data: \( D \mid O \), which is only logical given the distinction between "doers" and "non-doers." This is a central tenant of exposure modeling, one that is not universally recognized.

Response: We thank the reviewer for his comments and acknowledgement of this paper’s contribution.

3. The definition of ICC that is used should be explicitly presented when the term first appears. Currently, there is a "word definition" of the term, and it does not appear until p. 19! (The term is not commonly known, and there actually are alternative definitions in the literature.)
Response: We actually had the definition of ICC the first time we use it in the method section on Page 8:
“Intraclass correlation coefficients (ICC), the ratio of between-subject variance to the sum of between- and within-subject variance, were calculated as indicators for consistency of an individual’s activities.”

4. Correlation coefficients should be designated with an "r" rather than "R."

Response: This has been corrected.

5. The Discussion does not provide year of publication for the studies cited, and is inconsistent with prior citation style.

Response: This has been corrected.

6. I am very surprised to not see that age and gender do not have a major impact on any of the time use statistics presented in the tables. It is our experience that those two factors make the biggest impact on time use, much greater than day-type or season. Were those two variables not analyzed? If not they should be, as many of the other findings are conditional on what variables were included or not in the analyses.

Response: In the analysis of time-location data, we did not include age and gender in the analysis of diary data, partly because of the age and gender distribution of our study population. We broke down the analysis by age group, and the age of people in the same group were distributed across relatively small ranges. Moreover, as shown in the demographic table, the respondents in the “parents of young children” group were predominantly female, which is also part of the reason why we did not include gender as a covariate.

However, we agree with the reviewer that age and gender should be included in the analysis, as they have shown significant impact on time allocation in previous studies. We have re-run the analysis and updated Table 2 and the corresponding results and discussion.

Update in the “Results” section on Pg 12:

“Variation by demographic factors

Sex – In general, fathers of young children spent more time at work, shopping, eating and running errands than mothers; however, when all adults were considered, females were more likely to visit a food store than males. Male doers spent more time in other locations, including public parks and health clubs. In addition, girls who went to school or daycare spent more time at school than boys.

Age – Age effect was also examined within each age group. Older parents of young children were less likely to spend the whole day at home and more likely to spend time in transit and other locations, like parks and museums, than younger parents of young children. Time sleeping decreased with age for both young children and parents of young children. Among the older adult
doers, time spent at work, shopping, eating and running errands decreased with age. Older adults generally spent less time working than adults who were younger and had children.”

Update in the “Discussion” section on Pg 20:
“Graham and McCurdy considered age and gender as the primary factors to define a cohort in a time-activity study [21]. We did observe statistically significant impacts of sex, age and employment status on time spent in some locations and on some activities. However, our younger adult population comprised solely of parents of young children, which may have influenced their time activity patterns. Therefore, the variation by sex and age we observed may not be generalizable to populations of different characteristics (employment, family size).”

7. The comment that “sleep” as an activity is about 50% of the time spent at home, but I get approximate estimates of 59% in children, 36% in parents, and 46% in older adults. More heterogeneity than first implied.

Response: We thank the reviewer for the comments. Based on our rough estimate, parents spent 49% of their time at home sleeping, which is roughly 50%. We agree that children sleep more than adults. We note that the reviewer probably got the 36% value mistakenly by using the average time spent on work by parents of young children (397 min), which is right below the mean of time spent sleeping in by parents of young children Table 1, divided by average time spent at home (1101 min). We have revised the statement to

“..., and about half of the time at home was reported as sleeping; children slept longer than adults.”

8. There are a number of undefined terms appearing in the tables, and some symbols are used differently in different tables (for instance “P” for percentile in Table 1 and for “younger adults who were parents of young children” in Table 2. Other defined terms in Table 1 are “SD” and “Med.” While pretty obvious what they mean, they should be defined at first usage, at least. Also confusing to some may be the two meaning of “O” in Table 2.

Response: We thank the reviewer for noting that. We changed the “P90” in Table 1 and Table 3 to “90th%”. We have made the footnotes for “SD” and “Med” in Table 1. To avoid the confusion of two meanings of “O” in Table 2, we changed the abbr. for the three age groups to “children”, “parents”, and “older”.

9. The actual “r” and its “p” value should be shown in the last column of Table 2 instead of the qualifiers used. The footnotes in that should be cleaned up: for instance what does the second sentence in footnote “e” mean?

Response: We thank the reviewer for his comments. We have replaced the qualifier with the quantitative correlation measure. The results for time spent in residence / at own home were adjusted, and only the result for time spent at own home was kept, so the confusion in previous footnote ‘e’ (current footnote ‘f’) was cleared.
10. I would use “food store” rather than “grocery” when that land use is discussed.

    Response: This change has been made through the manuscript.

11. The moderate/vigorous data depicted in Table 3 are extremely high when compared with “objective monitoring” of physical activity (PA) that is seen in the literature (hundred of articles/year on this subject). The data are even much higher than seen in the PA survey, known to over-estimate participation and duration rates. In fact, the information presented on PA in Table 3 is not credible. One of the reasons for this is the rather loose definition of what constitutes moderate and vigorous PA (MPA & VPA). Some of the example activities provided on p. 13 are considered to be “light” PA by most exercise physiologists. Their definitions revolve around the METS concept, and a METS of ~ 3-6 usually is designated as MPA; >6 is considered to be VPA. See Ainsworth et al. (1993) Compendium in Med. Sci. Sports Exer. (I can provide a full citation if needed, but the above should do it.) Standing, for instance, is a light activity, with a METS <2.2. I do not have time to go into this subject in any detail, but the authors should do a fast “Google Scholar” search on MPA and VPA and review some of the articles that will appear. Both objective and subjective (survey) estimates should be briefly reviewed, to get an understanding of how much over-estimation of PA occurs in a survey, even in those who carry “activity logs” with them.

    Response: We thank the reviewer for the comments and valuable information. We have added the following explanation about our definition of moderate and vigorous PA into the results and discussion sections.

    In “Results” section on Pg 14:

    “Moderate activities were defined in the questionnaire by providing examples of walking, cleaning, food preparation, and other activities generally requiring standing up. Examples for vigorous activities in the questionnaire were running, bicycling, digging, and building. The “moderate” and “vigorous” activities were chosen such that they would be easily understood by respondents, but do not necessarily correspond to strict definitions based on available metabolic equivalent (MET) intensity levels for physical activities (Ainsworth et al., 2000). Thus compared to stricter definitions of moderate and vigorous physical activities, we may over-estimate the likelihood and duration of such behaviors.”

    In “Discussion” section on Pg 23:

    “Secondly, our definitions of moderate and vigorous activities were not precise, and tend to over-estimate the activity level. According to the activity metabolic equivalent (MET) intensity levels published by Ainsworth et al. [22], activities with METs ranging 3-6 were considered moderate and those with METs above 6 were considered vigorous. For example, walking can be light to vigorous depending on the speed (METs ranging 2.0 to 12.0), and bicycling varies from moderate to vigorous with METs of 4.0 to 16.0. Food preparation is more commonly considered to be a light activity with METs between 2.0 and 3.0, and digging is considered moderate with a MET of 5.0. Comparisons with available data on strictly-defined moderate and vigorous physical activities (MPA & VPA) also suggest overestimation of time spent on these activities. According to the State Indicator Report on Physical Activity [23], 67% and 45% of adults in California are physically active and highly active,
respectively. Ainsworth et al. [22] reported that adults in their 40’s spent 16 min/day in MPA and 18 min/day in VPA. Sallis et al. [24] investigated moderate and vigorous physical activities among 2126 adults between 20 and 74 years old. They found that 84% of respondents engaged in moderate activities and spent 50-83 minutes/day on average depending on respondents’ age; 15% of respondents reported vigorous activities and spent 17-68 minutes/day on vigorous activities. Compared to their study, we over-estimated the percentages of doers of moderate and vigorous activities by around 7% and 12%, respectively, and over-estimated the time spent on moderate and vigorous activities by approximately 3 times.”

12. Appendix I. I take it that none of the demographic variables that were obtained affected the D I O results? That frankly is hard to believe (see above comment regarding age and gender).

Response: As the reviewer suggested, we have included age and gender into the analysis and updated the results and discussion as presented in the answer to comment 6.

13. Example of a paragraph containing awkward sentence structure and a rewrite of same.

SUPERB enrolled 626 households in a telephone survey to assess time-activity, food recall, and the use of consumer products [13]. From these households, we recruited 250 for self-administered web surveys that covered the same subject matter as the telephone interview, of which 206 completed part of or all of the time-activity surveys. Only subjects able to complete the survey in English were recruited. Data were collected from two sub-cohorts, the first included 186 households in northern California in which at least one child under 8 years of age lived, and we enrolled both one parent and one child (not necessarily the younger one) from each household. Participants were selected from birth certificate records of children born between 2000 and 2005. The second cohort included 64 households of older adults (mostly ≥55-year-old) living in the southern portion of the Central Valley of California. These households were randomly selected from all tax-assessor’s housing units.

SUPERB enrolled 626 households in a telephone survey to assess the resident’s time use, eating habits, and consumer products usage [13]. From this sample, a subset of 250 households were chosen to complete a self-administered web survey that covered the same subject matter as the telephone interview; 206 households of this subset completed part of or all of the time-activity survey. Only subjects able to complete the survey in English were recruited. Data were collected from two sub-cohorts. The first included 186 households in northern California having at least one child > 8 years of age, and we enrolled one parent and one child (not necessarily the youngest one) from each household. These participants were selected from birth certificate records of children born between 2000 and 2005. The second cohort included 64 households of older adults (mostly ≥55-year-old) living in the southern part of California’s Central Valley. These households were randomly selected from housing units listed in County tax assessor records.

Response: We thank the reviewer for the comments and the revision. We have added the reviewer’s revised paragraph in the manuscript.

Examples of other grammatical/sentence structure problems (a lot of run-on sentences)
Response: We thank the reviewer for the comments. The revised sentences were shown below the original sentences.

A small number of participants (N=20) joined the study too late to complete 18-months of surveys completing only 15-months.

A small number of participants (N=20) joined the study too late to complete the full 18-months of surveys and completing only 15-months of surveys.

The performance and reliability of the web surveys collected were evaluated, e.g., the survey completion rate, the percent of surveys with unreasonable time reported as spent sleeping.

The performance and reliability of the web surveys collected were evaluated in terms of, e.g., for example, the survey completion rate, and the percent of surveys with an unreasonable amount of time reported as spent sleeping.

Several factors appeared to influence the time spent in selected microenvironments and activities varied (Table 2).

Several factors appeared to influence the time spent in selected microenvironments and activities varied (Table 2).

Higher ICCs were observed for time that children spent in the school and adults spent in the office and on working, probably because school activity and working activity were kind of routines and had less variability within an individual.

Higher ICCs were observed for time period that children spent in the school and adults spent in the office and on working, probably because school activity and working activity were kind of routines and had less variability within an individual.

Another interesting longitudinal trend is that, with time passing, children were more likely to use a computer, but less likely to watch TV.

Another interesting longitudinal trend is that, with as time passing progressed, children were more likely to use a computer, but less likely to watch TV.

The ICCs for time spent on selected activities were generally moderate, ranging from 0.14 to 0.58 for activity that happened on a daily basis and even higher (0.30-0.73) for the frequency of activities that happened less often, suggesting variability in activity pattern between individuals.

The ICCs for time spent on selected activities were generally moderate, ranging from 0.14 to 0.58 for activity that happened on a daily basis, and even ICCs were higher (0.30-0.73) for the frequency of activities that happened less often, suggesting variability in activity pattern between individuals.
In addition, previous studies examining variations in time-location/activity patterns were based on very specific populations or were more restricted with regard to locations they included, our study provides additional evidence that helps to confirm and consolidate previous findings.

In addition, previous studies examining variations in time-location/activity patterns were based on very specific populations or were more restricted with regard to locations they included. Our study provides additional evidence that helps to confirm and consolidate previous findings.

A limitation of this study is that, during the study, participants gradually withdrew over time, which hinders the evaluation of longitudinal variation to some degree.

A limitation of this study is that, during the study, participants gradually withdrew over time, which hinders the evaluation of longitudinal variation to some degree.

In addition, web surveys allow participants to select “convenience” day for recall, however may increase the possibility that participants select unrepresentative days with less activities than normal to minimize reporting effort, for example, a nurse could select a weekday that he/she was not on shift for recall.’

In addition, web surveys allow participants to select any day for recall, and they may select a convenient day. Specifically, “convenience” day for recall, however may increase the possibility that participants may select an unrepresentative days with less activities than normal to minimize reporting effort, for example, a nurse could select a weekday that he/she was not on shift for recall.

Information on participation in moderate and vigorous activities: there are hundreds of these articles published every year. The following were within arm’s reach when I wrote this.

CDC (2009). Early release of selected estimates based on data from the January-March 2009 NHIS (p. 46). Can obtain from the web. Percentage of adults engaging in regular leisure-time physical activity (Figure 7.1): generally in the 30-35% range.

CDC (2010). State Indicator Report on Physical Activity, 2010. Percent of adults in California that are physically active (66.7%) and highly active (45.0%). www.cdc.gov/physicalactivity/.

Lanctot et al. (2008). Participation of obese African-American teens in moderate (MPA) and vigorous (VPA), or both (MVPA): 3 min/day.

Ainsworth et al. (2000). Adults in their 40’s: 16 min/day in MPA; 18 min/day in VPA; 35 min/day in both.

Behrens et al. (2005). Young adults: 28 min/day in MPA; 4 min/day in VPA.

Response: We thank the reviewer and have included these studies.
This concludes my review.
Reviewer 2

Title: Longitudinal Variability of Time-Location/Activity Patterns of Population at Different Ages

Version: 1 Date: 29 April 2011

Reviewer: Kai Elgethun

Reviewer's report:

When assessing the work, please consider the following points:

1. Is the question posed by the authors new and well defined? New: NO. Well-defined: YES.

2. Are the methods appropriate and well described, and are sufficient details provided to replicate the work? YES

3. Are the data sound and well controlled? YES

4. Does the manuscript adhere to the relevant standards for reporting and data deposition? YES

5. Are the discussion and conclusions well balanced and adequately supported by the data? NEEDS WORK—SEE BELOW

6. Do the title and abstract accurately convey what has been found? YES

7. Is the writing acceptable? YES

- Major Compulsory Revisions (which the author must respond to before a decision on publication can be reached)

Validity and Reliability: the authors must point out that a MAJOR WEAKNESS of this study is that there are no measures of reliability/concordance or validity for respondent’s answers. I recommend several sentences in the Methods and several sentences in the Discussion to point out why this is a major weakness. This should also be stated in the Abstract.

This was also a weakness in first published article from the SUPERB study (Environmental Health 2010, 9:54)

Response: We thank the reviewer for the comments. We have addressed the weakness in the Method and Discussion section as presented below.
‘Reliability’ is mentioned in the Methods under ‘Data Analysis’ as referring to a way that the 3% substandard quality surveys were excluded. This is not reliability. This is simply quality control. Please be clear in your nomenclature. Reliability has a specific meaning.

Response: We agree with the reviewer. It is not appropriate to refer these quality control steps as “reliability”. We have reworded these sentences, as shown in the responses to the next comment.

From Will G. Hopkins, 2000
(http://www.sportsci.org/resource/stats/precision.html):

The two most important aspects of precision are reliability and validity. Reliability refers to the reproducibility of a measurement. You quantify reliability simply by taking several measurements on the same subjects. Poor reliability degrades the precision of a single measurement and reduces your ability to track changes in measurements in the clinic or in experimental studies. Validity refers to the agreement between the value of a measurement and its true value. You quantify validity by comparing your measurements with values that are as close to the true values as possible. Poor validity also degrades the precision of a single measurement, and it reduces your ability to characterize relationships between variables in descriptive studies.

Please refer to http://www.socialresearchmethods.net/kb/reltypes.ph

For a brief primer on the types of reliability. Such tests of reliability can be a question that re-states an earlier question, or a comparison to some other measure.

In the case of time-location, self-report diary data should really be checked against some other measure of time-location. Discrepancies between measures of time-location should be quantified whenever possible and the difference reported as an error range/confidence interval or Cohen’s Kappa statistic. Inter-rater tests are well-established in the social science and social epidemiology literature. Refer to “Principles of Exposure Measurement in Epidemiology”, Ch.4, Validity and Reliability Studies (Armstrong, White and Saracci, 1994). Many technologies exist to use as a comparison to a diary, and have been used for at least a decade for this purpose.

Response: We thank the reviewer for the comments. We actually have conducted extensive evaluation on the reliability and validity of the web survey we used. The method evaluation paper is currently in Press with the Journal of Exposure Science and Environmental Epidemiology. We regret that we did not provide this other paper to the reviewers and have included it now.
Regarding reliability, it is not very practical to collect several time diaries on the same subject for one time period, so the alternative option is to collect multiple diaries at different time points on the same subjects, which is what we did. However, the multiple diaries we collected at different time points from the same subjects were purposely used to evaluate the longitudinal variation of human activity pattern, namely, we expect variability due to time variables, such as, day-type, season, and wave, between the diaries from the same subjects. In other words, the longitudinal variation we are evaluating is part of and cannot be clearly separated from the real “reliability”, as the real “reliability” is attributed to respondents’ compliance in completing the surveys in a consistent way. Thus, we could not test the reliability based on the reproducibility of the diaries we collected. Instead, we tested the reliability by evaluating the consistency of number of records and number of location change reported in diaries from the same subject using a Generalized Liner Mixed Model, with day-type (weekday vs. weekend), season (warm season from May to October vs. cool season from November to April), and wave number or type of survey interface as fixed effect estimates and individual participants as random effect estimates. Such variation also indicates changes in participant compliance over time due to a loss of enthusiasm with time. No significant variation over time or between the two web survey interfaces was observed, suggesting consistent compliance and reproducibility over the study period as long as subjects did not decide to completely drop out of the study. These results are presented in the method evaluation paper.

We also evaluated the validity by comparing the web survey results with data collected by telephone interviews from the same respondents, which is the traditional method for collecting time-activity data with interviewers’ guidance. However, the web survey and telephone interview were not conducted on the same day for the same respondent, with telephone survey 0.65 year earlier on average than the first wave of web survey from the same subjects, as it would be too burdensome by asking respondents to make two recalls for the same day. Therefore, we presented the distribution of the data collected by web survey and telephone interview respectively, and conducted pair-wise comparison of time-location data in the first wave of the web survey and telephone survey from the same subjects. Results showed that respondents’ time-location data between two methods were significantly correlated on weekdays but not significantly correlated on weekend days, which is sensible since people probably had different weekend plans in different seasons. Given such results, we did not further calculate CIs or Cohen’s Kappa. We agree that comparison with some objective method, such as GPS data, can better determine the validity of the web survey, but, unfortunately, it was not in our study plan.

We do agree with the reviewer that such method reliability information should be included in this manuscript. As the reviewer suggested, we added more method evaluation results and addressed the limitation in the abstract, method and discussion section, as follows.

In the method section on Pg 8:
“A full evaluation of the web-survey method and data cleaning have been described previously in Wu et al. [17]. Reliability was evaluated by examining the consistency of the number of records and number of location changes reported in a diary, and the validity was tested by comparing with data collected through telephone interviews of the same respondent. Both the reliability and validity were acceptable for this web survey method. The performance of the web surveys collected was
evaluated in terms of, for example, the survey completion rate and the percent of surveys with an unreasonable amount of time reported for sleeping. We observed compliance issues in a small number (<3%) of web diaries, and these diaries were not included in the analysis. Diaries were considered sufficiently complete for inclusion: if (a) they contained three or more time-location-activity records; and (b) information was missing for less than 3-hours per day.”

In the discussion section, at the end of the limitation paragraph on Pg 24:
“Lastly, in our method evaluation step, we compared the web survey data with the time-activity data collected by telephone interview from the same respondents and obtained similar distributions of the time-allocation measured by these two methods [17]. However, since these two types of the surveys were not conducted for the same day for a respondent, thus such comparison does not fully establish the validity of the web survey method. More objective methods, such as Global Position System (GPS) recording, may provide better reference values to determine the validity of the self-reported location data [25]; gold standards for activity data are more difficult to obtain.”

Your group mentions GPS data collection in the first published article from the SUPERB study (Environmental Health 2010, 9:54)

What happened to these time-location data? It seems like these data belong here in a discussion of time-location patterns. If the GPS data are to be analyzed and presented later, please give details. If not, why?

Response: We thank the reviewer for the thorough review of our previous publication. However, the GPS data were collected in a different subsection of the overall study. They were collected from different respondents from those participated in the web survey, so we were not able to conduct a parallel comparison between the web survey data and GPS data. We hope to analyze and present these data in a future paper.

This manuscript MUST mention the relevant literature on this issue to be considered complete and publishable. Search “time location diary” in PubMed.

If the cohort is too big to test validity and reliability for everyone, a sub-cohort can be tested and a measure of precision reported.

In addition to the deficiencies mentioned above, the References should be revisited completely, as there have been several new time-use articles published in the last 1-2 years that were not referenced in this manuscript. This can often happen if a manuscript’s introduction is written a few years before the manuscript is submitted, I realize.

Response: We thank the reviewer for the comments. We have cited following recent publications on time-activity variability.


- Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

  Response: This has been corrected.

- Discretionary Revisions (which are recommendations for improvement but which the author can choose to ignore)

Reference 12 and 13 are the same.

  Response: This has been corrected.

Otherwise OK. The writing is good and the analyses of data collected are sufficient. This is a good use of web-based surveys. The cohort is sizeable.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**
I Declare that I have no competing interests