Implications of Measurement: Comparing Estimates of Physical Activity across NHANES, NHIS, and ATUS

Rachelle Hill[1], Sarah Flood[2] and Kari Williams[2]

[1]U.S. Census Bureau

[2]University of Minnesota

**Any opinions and conclusions expressed herein are those of the authors and do not necessarily represent the views of the U.S. Census Bureau or the University of Minnesota. All results have been reviewed to ensure that no confidential information is disclosed.

Prior research has clearly demonstrated the importance of physical activities for health and, as such, the literature investigating such patterns is diverse. The diversity of the literature is partially a reflection of the numerous data sources available and the many ways to measure physical activity. Understanding how estimates of physical activity and relationships between covariates and physical activity differ by physical activity measurement is thus an important contribution to the limited research base about physical activity measurement. And, investigations of the correlates of physical activity have important implications for the health of the population because it is a potentially modifiable health behavior. The various data and approaches to measurement we consider here -- self-reported measures of frequency and intensity, time diary data, and accelerometer data -- each have their merits. Despite its importance for health, scholars are keenly aware of the potential inaccuracies of self-report data (inaccurate and biased reporting) and time diary data (limited visibility into usual physical activity given the typical one-day diary design). Even accelerometer data, the gold standard for measuring physical activity, is not without its challenges (such as participant burden). Using nationally representative data from the 2003-2006 National Health and Nutrition Examination Survey (NHANES), the 2003-2013 National Health Interview Survey (NHIS), and the 2003-2013 American Time Use Survey (ATUS), we propose to investigate the following questions. First, how do recall estimates of time spent in physical activity in the ATUS and NHIS compare to NHANES estimates based on accelerometer data? Second, how do ATUS estimates of time spent in activities above different metabolic equivalent (MET) value thresholds compare to NHANES estimates based on accelerometer data? And finally, to what extent do the measurement differences in these estimates vary by BMI and self-reported health?

Background

The duration, intensity, and frequency of physical activity is often discussed in scholarly and popular media as an important indicator of current and future physical health in part because exercise has been linked to decreased mortality, risk of heart disease, hypertension, colon and breast cancer, diabetes, and depression (U. S. Department of Health and Human Services 1996; U.S. Department of Health and Human Services 2010). Despite the established associations between physical activity and health outcomes and the high number of studies investigating physical activity, measuring physical activity is challenging and concerns regarding best practices remain.

Accelerometer data is often considered to be the “gold standard” of physical activity measurement. Accelerometers, in particular those used by NHANES participants, are small devices worn on the hip of a study participant that track the duration and intensity of common activities used for locomotion (e.g., walking or running). Research has shown that time spent in physical activity (especially when limited to bouts of 10 minutes or more) is low, peaking in childhood and quickly declining across the life course for both men and women as well as all racial and ethnic groups (Troiano et al. 2008), with few adults meeting the minimum recommendations for physical activity (Tucker, Welk, and Beyler 2011). Though the data is considered to be highly objective and often more complete than self-report data, there are several important barriers to using accelerometer data including the cost of devices, the burden to participants, and the interpretability of data. In particular, the cost and logistical challenges inherent in maintaining the devices make it nearly impossible to assess physical activity across a nationally representative sample. Such challenges make other measurements of physical activity, such as self-report or time diary data, highly attractive.

Self-reported data relies on the recollections of respondents about their physical activity. Self-reports, typically collected through interviews, diaries/logs, or questionnaires, are typically low cost and low burden, which can facilitate larger and more representative samples. Recall and response bias represent potential challenges to reliable data as participants may over- or under-report their physical activity because of social desirability, inaccurate memories, or lack of clarity about which physical activities are to be included (Adams et al. 2005; Prince et al. 2008). Additionally, while self-reports may inquire about the specific activity as well as its duration or intensity, self-reporting is not a direct measure and cannot capture the level of physical activity of the respondent. Prior research has demonstrated important differences in time in physical activity using self-report and accelerometer data (Tucker et al. 2011). The NHIS utilizes self-reported data through interview questions about the frequency, duration, and intensity of physical activity.

Time diaries are an alternative method of measuring all types of activities that individual's participate in during a set period of time. The ATUS is a time diary survey that asks individuals to describe all of the activities they participated in during the previous day starting at 4:00 AM the morning prior to the interview day until 3:59 AM of the interview day. One strength of the time diary method is that individuals are not asked about a specific type of activity and due to the breadth of activities participated in during a given day, individuals are less likely to be biased towards a specific activity (Juster and Stafford 1985). However, it is unclear how time diary estimates may compare to both self-report and accelerometer estimates of physical activity as we were unable to identify prior research assessing potential differences.

To develop a greater understanding of the different approaches to measuring physical activity, we investigate the following three research questions by drawing on nationally representative data from NHANES, NHIS, and ATUS. First, how do recall estimates of time spent in physical activity in the ATUS and NHIS compare to NHANES estimates based on accelerometer data? Second, how do ATUS estimates of time spent in activities above different metabolic equivalent (MET) value thresholds compare to NHANES estimates based on accelerometer data? And finally, to what extent do the measurement differences in these estimates vary by BMI and self-reported health?

Data and Methods

We draw on data from NHANES, NHIS, and ATUS to perform our comparisons of physical activity. For comparability across samples and due to our focus on working aged adults, we limit our analytic sample to respondents aged 18 to 64 with no missing data on physical activity.

NHANES

We draw on data from the 2003 to 2006 National Health and Nutrition Examination Survey (NHANES) (Centers for Disease Control and Prevention (CDC) and National Center for Health Statistics (NCHS) 2013). NHANES is a national study of health status, health behaviors and nutrition among a nationally representative sample of children and adults in the United States conducted by the National Center for Health Statistics (NCHS). Data were collected via household interviews and physical examinations completed in mobile examination centers. The full NHANES sample included 20,470 respondents and our analytic sample includes 4,641 individuals.

NHIS

We use the Integrated Health Interview Series (IHIS), a harmonized version of the NHIS dataset (Minnesota Population Center and State Health Access Data Assistance Center 2012). The NHIS has been fielded annually since 1957; since 1960 the survey has been conducted by the NCHS, which is part of the CDC. The NHIS is nationally representative of the non-institutionalized civilian population residing in the United States; this excludes patients in long-term care facilities, incarcerated persons, persons serving active military duty, and U.S. nationals living outside of the country. The sampling plan is redesigned after every decennial census and oversamples Black, Hispanic, and Asian persons. Data for the NHIS are collected through in-person household interviews conducted by trained employees of the U.S. Bureau of the Census. Since 1997 the interviewer uses a computer assisted personal interviewing (CAPI) mode, where trained interviewers use a laptop and enter responses directly into the computer during the interview. The NHIS contains a core questionnaire, the content of which changes little from year to year and allow for trends analysis and pooling data across years, as well as annual supplements used to respond to emerging and changing public health trends. For each family in the NHIS, one adult and one child (if any children are present) are randomly selected to answer additional questions through the Sample Adult Core and Sample Child Core; these collect information on health status, health care services, and health behaviors. The IHIS sample from 2003-2013 included 1,003,921 respondents, 248,911 of whom are 18-64 and contain full information on our independent variables.

ATUS

We use integrated American Time Use Survey (ATUS) data for our analyses (Hofferth, Flood, and Sobek 2013). The ATUS is a time diary study of a nationally representative sample of Americans and has been collected annually since 2003. ATUS sample members are invited to complete the survey following their exit from the Current Population Survey (CPS), which is a household survey of the civilian, non-institutionalized population. One individual aged 15 or older per former CPS participating household is randomly selected to respond to the ATUS during the two to five months following their exit from the CPS. ATUS respondents report the activities they engaged in over a 24-hour period from 4:00 a.m. of yesterday until 4:00 a.m. of the reporting day (diary day), as well as where, when, and with whom activities were done. Over 400 detailed activity codes are represented in the three-tier six-digit activity coding scheme. Data are collected on all days of the week, and weekends are oversampled. Though the data may not typify respondents' daily activities, aggregations of the data are representative of the American population. The ATUS sample from 2003-2013 included 148,345 respondents, 113,105 of whom are 18-64 and contain full information on our independent variables.

Measures

Physical Activity Measures

Our first physical activity measure indicates the total number of minutes respondents spend per day in episodes of physical activity. For comparability reasons, episodes of physical activity are only included if they are 10 minutes or longer in duration.

Our second physical activity measure is only available for NHANES and ATUS and indicates the total minutes per day respondents spend in physical activity above a given MET value thresholds of 3 (physical activity) and 6 (vigorous physical activity) (Tudor-Locke et al. 2009).

A key advantage to the NHANES is Mobile Exam Center where respondents go to participate in the health examination. During the exam, individuals are recruited to wear the physical activity monitor (Actigraph). Those younger than six and those subjects in wheelchairs or with other physical impairments that prevent walking are excluded from this portion of the study. The Actigraph is an accelerometer calibrated to measure physical activity, which is meant to objectively measure their physical activity levels (Centers for Disease Control and Prevention 2005). The activity monitor captured the intensity of activity at 1-minute intervals. Participants were asked to wear the Acitigraph on their right hip during their waking hours for seven days. The physical activity measure was constructed from the raw data collected by the Actigraph and included time spent in vigorous and/or moderate physical activity for episodes longer than 10 minutes. The raw data from NHANES was first edited using programs available by the Applied Research Cancer and Population Sciences division of the National Cancer Institute (2013) to make comparable to other datasets as well as to reweight the accelerometer data to match the larger NHANES examined sample. Time spent in moderate or vigorous activities is a continuous measure of time averaged for a single day. To maintain comparability across datasets only bouts of physical activity that were 10 or more minutes in duration (i.e., episodes of 9 or fewer continuous minutes were not included in this measure) were included in the measure.

Beginning in 1997, NHIS began including an inquiry about adult leisure-time physical activity in the Sample Adult Core, distinguishing between "vigorous" and "light or moderate" activity. Persons were asked a screener question about how often they engage in leisure time activities for at least 10 minutes that cause "heavy sweating or large increases in breathing or heart rate" for vigorous activity, and "only light sweating or a slight to moderate increase in breathing or heart rate" for moderate activity. Subsequent questions for persons who reported in engaging in such activities inquire about the frequency and duration of physical activity sessions for vigorous and moderate physical activities respectively. These can be combined to generate the average daily minutes of physical activity. Eligible respondents who reported that they do not engage in vigorous or moderate physical activity for bouts of 10 minutes or more are assigned values of zero for average daily minutes of physical activity in our analyses.

Our ATUS physical activity measure is derived from time diary reports about what respondents were doing and how they were doing it. Each episode of activity is recorded separately and the minutes spent in that activity are recorded. We consider diary episodes of at least ten minutes in duration for comparability with NHIS measures. The minutes spent in the episode count toward our physical activity measure if the respondent was exercising or playing sports (hunting and fishing are excluded) *or* if the respondent was walking or biking as a mode of transportation. Importantly, because the time diary only covers one 24-hour period, there will be zeros for several respondents since physical activity may not be a daily activity for many.

Key Health Measures

Our two key measures of health include *self-reported health*, distinguishing between excellent, very good, good, fair, and poor, and *body mass index (BMI)* with categories underweight (<18.5), normal weight (18.5-24.9), overweight (25-29.9), and obese (30+).

Other Measures

Gender is coded as male and female. *Age* of the respondent is measured as a categorical variable: 19-24, 25-34, 35-44, 45-54, and 55-64. *Race* is coded as four dichotomous variables: White, non-Hispanic (reference); Black, non-Hispanic; Other, non-Hispanic; and Hispanic. *Immigrant* indicates whether the respondent was native (reference) or foreign born. *Marital status* differentiates between married (reference), cohabiting, widowed/divorced/separated, and never married. *Education* is coded into four dichotomous variables: less than high school (reference in regression analyses), high school degree, some college, and college degree or more. *Income* is coded into three broad categories: up to $34,999, $35,000-$74,999 (reference), and $75,000 and higher. *Employment status* differentiates into full time (35+ hours per week), part time (1-34 hours per week), non-employed, and self-employed.

Analytic Method

Our analytic approach is a multi-step approach that aims to examine the sample composition, estimated averages of time spent in physical activity across the population as well as within subsamples, and by BMI and self-reported health. We will begin by comparing demographic characteristics across the weighted samples. Next, we will compare weighted estimates of time spent in moderate or vigorous physical activity as well as subsample estimates by gender and age. Finally, we will compare time spent in moderate or vigorous physical activity across categories of self-reported health and BMI. We begin by comparing the samples below. Next steps will include comparing stratified weighted estimates and estimating multivariate models and comparing those factors that are important predictors of time spent in physical activity, in particular demographic characteristics, self-reported health, and BMI.

Preliminary Results

In Table 1 we compare the sample composition of NHANES, NHIS, and ATUS. Using the sample weights, each sample is meant to be nationally representative and should have similar sample characteristics. Here we see that approximately 50.9%, 50.9% and 50.7% of the sample is female in NHANES, NHIS, and ATUS respectively. Though there is some variation across demographic characteristics, there are also similar proportions of each race/ethnicity, immigrant status. There is less variation than expected in BMI categories (e.g. 32.8%, 36.3%, and 34.6% are normal weight in NHANES, NHIS, and ATUS respectively) especially considering that weight and height are self-reported measures in NHIS and ATUS and are measured during the exam in NHANES. There is more variation across the samples than expected in characteristics like employment status, education, and income potentially reflecting non-response due to respondent burden. There is also some interesting variation in self-reported health as the proportion of NHANES respondents that report being in excellent health is notably smaller than NHIS or ATUS (13.1% of NHANES respondents compared to 31.4% of NHIS respondents and 20.4% of ATUS respondents). Finally, we see important differences in minutes spent in average daily physical activity. The mean time spent in physical activity (including bouts of 10 minutes or more only) is 8.8, 31.7, and 18.8 minutes per day for NHANES, NHIS, and ATUS respectively. The differences are as would be expected with NHIS having the largest mean minutes of daily physical activity perhaps due to overestimation of time spent exercising or difficulty recalling and aggregating time spent in physical activity. NHANES estimates are lowest, which may be surprising because accelerometer data are collected continuously with the Actigraph, but may reflect the limited time spent in physical activity as well as respondent burden or cooperation in terms of having to wear the Actigraph to accurately measure physical activity. ATUS estimates fall between the NHANES and NHIS estimates. Based on what

we know from literature comparing time diary and stylized questions about time spent in paid work and housework (e.g., Juster, Ono, and Stafford 2003) and reading to children (e.g., Hofferth 2006), it is not surprising that ATUS estimates of physical activity are lower than NHIS estimates.

Conclusion

Physical activity is an important indicator of current and future health for individuals and scholars have invested a great deal of time and energy in understanding not only who engages in physical activity but also the outcomes associated with different levels of physical activity. Despite its importance, physical activity is difficult to measure. To further our understanding of the methodological challenges associated with physical activity we attempt to compare three different techniques for measuring physical activity using three different nationally representative samples. Preliminary results sought to illustrate similarities and differences across NHANES, NHIS, and ATUS and we see that the samples are in many ways very similar including gender and racial/ethnic composition. However, notable differences across the distribution of income, self-reported health, and physical activity raise interesting questions and will direct our future investigations. Next steps will investigate these differences further while also examining differences in time spent in physical activity using multivariate methods.

Table 1. Comparison of Demographic Characteristics by Survey

	NHANES 2003-2006	NHIS 2003-2013	ATUS 2003-2013	
Daily Physical Activity (mean minutes)	8.89	31.71	18.84	[123]
BMI				
Under weight (<=18.5)	1.7	1.21	1.73	[13]
Normal weight (18.5-24.9)	32.81	36.28	34.57	[13]
Overweight (25-29.9)	32.74	35.99	35.81	[12]
Obese (30 or greater)	32.76	26.51	27.9	[123]
Self-Reported Health				
Excellent	13.15	31.41	20.45	[123]
Very Good	37.62	33.09	35.09	[123]
Good	36.08	24.81	30.07	[123]
Fair	11.43	8.13	11.23	[13]
Poor	1.73	2.57	3.17	[123]
Gender				
Women	50.95	50.97	50.73	
Men	49.05	49.03	49.27	
Age				
19-24	11.72	13.64	13.15	[123]
25-34	23.34	21.95	22.01	

Table 1 cont. Comparison of Demographic Characteristics by Survey

35-44	24.59	22.6	22.78	[12]
45-54	25.19	23.51	23.64	[1]
55-64	15.16	18.3	18.42	[12]
Race				
Non-Hispanic White	69.67	66.57	67.76	[123]
Non-Hispanic Black	12.11	11.94	11.71	
Hispanic	12.62	14.91	14.85	[12]
Other	5.6	6.58	5.69	[13]
Immigrant				
Born outside of the U.S.	16.32	17.46	16.63	[3]
Born in the U.S	83.68	82.54	83.37	[3]
Marital Status				
Married	59.37	56.34	57.85	[13]
Cohabiting	7.98	7.7	5.11	[23]
Widowed/Divorced/Separated	13.35	13.18	13.11	
Never Married	19.22	22.78	23.93	[123]
Educational Attainment				
Less than High School Degree	13.25	13.03	10.83	[23]
High School Degree or Equivalent	24.52	26.76	30.41	[123]
Some College or Associate's Degree	34.73	31.61	28.37	[123]
College Degree or Higher	27.49	28.6	30.39	[23]
Family Income				
Up to $34,999	34.15	32.29	31.1	[123]
$35,000 to $74,999	36.21	33.33	35.4	[13]
$75,000 and up	29.64	34.38	33.5	[123]
Employment status				
Part-time (less than 35 hours per week)	7.88	12.86	11.82	[123]
Full-time	64.56	53.41	56.07	[123]
Self-employed	7.49	7.77	7.74	
Not employed	20.08	25.95	24.38	[123]
N	4,641	248,911	113,105	

Notes: Means are weighted. [1]=NHANES mean significantly different than NHIS mean ($p<.05$); [2]=NHANES mean significantly different than ATUS mean ($p<.05$); [3]=NHIS mean significantly different than ATUS mean ($p<.05$).

References

Adams, Swann Arp et al. 2005. "The Effect of Social Desirability and Social Approval on Self-Reports of Physical Activity." *American Journal of Epidemiology* 161(4):389–98.

Applied Research Cancer Control Population Sciences. 2013. "SAS Programs for Analyzing NHANES 2003-2004 Accelerometer Data." Retrieved March 1, 2013 (http://appliedresearch.cancer.gov/nhanes_pam/).

Centers for Disease Control and Prevention. 2005. *Anthropomentry and Physical Activity Monitor Procedures Manual*. National Center for Health Statistics. Retrieved September 27, 2013 (http://www.cdc.gov/nchs/data/nhanes/nhanes_05_06/BM.pdf).

Centers for Disease Control and Prevention (CDC), and National Center for Health Statistics (NCHS). 2013. *National Health and Nutrition Examination Survey Data*. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention. Retrieved September 27, 2013 (http://www.cdc.gov/nchs/nhanes.htm).

Hofferth, Sandra L. 2006. "Response Bias in a Popular Indicator of Reading to Children." *Sociological Methodology* 36(1):301–15.

Hofferth, Sandra L., Sarah M. Flood, and Matthew Sobek. 2013. *American Time Use Survey Data Extract System: Version 2.4 [Machine-Readable Database]*. Maryland Population Research Center, University of Maryland, College Park, Maryland, and Minnesota Population Center, University of Minnesota, Minneapolis, Minnesota.

Juster, F. Thomas, Hiromi Ono, and Frank P. Stafford. 2003. "An Assessment of Alternative Measures of Time Use." *Sociological Methodology* 33:19–54.

Juster, F. Thomas, and Frank P. Stafford. 1985. *Time, Goods, and Well-Being*. Ann Arbor, MI: Institute for Social Research.

Minnesota Population Center and State Health Access Data Assistance Center. 2012. *Integrated Health Interview Series: Version 5.0*. Minneapolis, MN: University of Minnesota. Retrieved September 18, 2014 (http://www.ihis.us).

Prince, Stéphanie A. et al. 2008. "A Comparison of Direct Versus Self-Report Measures for Assessing Physical Activity in Adults: A Systematic Review." *International Journal of Behavioral Nutrition and Physical Activity* 5(1):56.

Troiano, Richard P. et al. 2008. "Physical Activity in the United States Measured by Accelerometer." *Medicine & Science in Sports & Exercise* 40(1):181–88.

Tucker, Jared M., Gregory J. Welk, and Nicholas K. Beyler. 2011. "Physical Activity in U.S. Adults: Compliance with the Physical Activity Guidelines for Americans." *American Journal of Preventive Medicine* 40(4):454–61.

Tudor-Locke, Catrine, Tracy L. Washington, Barbara E. Ainsworth, and Richard P. Troiano. 2009. “Linking the American Time Use Survey (ATUS) and the Compendium of Physical Activities: Methods and Rationale.” *Journal of Physical Activity and Health* 6(3):347–53.

U. S. Department of Health and Human Services. 1996. *Physical Activity and Health: A Report of the Surgeon General*. U.S. Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion.

U.S. Department of Health and Human Services. 2010. “Healthy People 2020 - Improving the Health of Americans.” *Healthy People 2020*. Retrieved December 29, 2010 (http://www.healthypeople.gov/2020/default.aspx).