Epi Research Report

New York City Department of Health and Mental Hygiene April 2013

Self-Reported and Accelerometer-Measured Physical Activity: A Comparison in New York City

In 2010-2011, the New York City Department of Health and Mental Hygiene conducted the Physical Activity and Transit (PAT) survey, a random-digit-dial telephone survey of adult New Yorkers that tracked levels of sedentary behaviors and physical activity at work, in the home, for recreation and as active transportation (e.g., walking and biking). The survey also addressed other questions about factors that facilitate active lifestyles. A subset of those who completed an interview also was asked to take part in a follow-up study in which their activity levels were measured objectively using an accelerometer device. This Epi Research Report examines the relationship between self-reported physical activity and accelerometer-measured physical activity and provides best practice recommendations for using self-reported data.

Physically active adults have lower rates of various chronic diseases and premature death than those who are less active.[1,2] Physical activity is difficult to measure in population-based health surveys. However, reliable and valid measurements are needed to further study the link between physical activity and health outcomes, to track population levels of physical activity through surveillance and to evaluate interventions to increase physical activity.[3] In 2010-2011, the New York City Department of Health and Mental Hygiene conducted the Physical Activity and Transit (PAT)

1 Physical Activity Guidelines Advisory Committee. Physical Activity Guidelines Advisory Committee Report, 2008. Washington, DC: U.S. Department of Health and Human Services, 2008.
2 Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. *CMAJ*. 2006; 174(6):801-809.
3 Mâsse LC, de Niet JE. Sources of Validity Evidence Needed With Self-Report. *J Phys Act Health*. 2012; 9 (Suppl 1):S44-55.

Key Findings:

- Physical Activity and Transit (PAT) survey participants reported higher levels of activity than were measured by accelerometers.
- Self-reported and accelerometer-measured minutes of physical activity were not normally distributed in the PAT data. Medians rather than means should be used to summarize the estimates of minutes of physical activity.
- The relationship between self-reported and accelerometer data differed by gender and age of the participant. Males underreported their physical activity, while females and older adults (65 years and older) overreported. Self-reported estimates should be stratified by gender and/or age when comparing groups.
- Self-reported data on physical activity measured by surveys cannot always accurately quantify absolute numbers of minutes. However, these data can provide details on the type, context and purpose of physical activity not captured by accelerometers, and effectively capture relative differences in activity between groups and over time.

NYC Health

What is an accelerometer?

Accelerometers are small electronic devices that measure movement. Unlike a pedometer, which measures steps taken by an individual, accelerometers measure the amount of movement and its level of intensity. Accelerometers provide objective data on the periods of time someone is physically active throughout the day and the intensity of the physical activity. Calibration studies compare "counts" – the units of measurement in accelerometry – to observed activities with established levels of energy expenditure. From these studies, cut-points were developed to process accelerometer data into intensity categories (e.g., light activity, moderate activity and vigorous activity).[4] Thus, intensity count values provide objective standards for measuring the intensity of activity throughout the day.

survey, which measured physical activity levels and duration with an accelerometer device.

Although self-reported data can provide information on the type, context and purpose of activity, which is unavailable through objective measurement, accelerometer data are considered to more accurately measure absolute levels of physical activity.[4] The study presented here used the accelerometer data as an objective point of reference to evaluate self-reported measurements of physical activity. This report provides recommendations for the best ways to use self-reported data on physical activity based on a comparison with a more objective accelerometer measure.

Methods

Collecting the data

The PAT survey was a random-digit-dial survey (including cell and landline telephones) conducted in two waves that when combined, provided data from 3,811 completed interviews.

The first wave of interviews was conducted between September and November of 2010 (n=1,323). The second wave was conducted between March 2011 and November 2011 and included 2,488 interviews, of which 1,261 were from four areas of the city with the highest prevalence of obesity (South Bronx/Harlem, East/Central Brooklyn, East Queens/Rockaways and Northern Staten Island).

Individuals interviewed in the second wave of the PAT survey were asked to participate in the device follow-up study using accelerometers to objectively record their activity levels. Participants were asked to wear the accelerometer for one week during all waking hours and to remove it only when in water (see protocol in Table 1). Participants in the device follow-up study also wore GPS devices to track location; these data were not included in the research presented in this report. Of those who completed an interview in the second wave of the PAT survey (n=2,488), 803 (32%) agreed to participate and returned devices with data. The minimum accelerometer wear-time for a reliable estimate of weekly activity is 10 or more hours on four or more days.[4] With this cut-off as the inclusion criterion, 679 participants were included in the final analytic dataset used for this study.

Measuring physical activity by self-report

Self-reported physical activity levels were measured using the Global Physical Activity Questionnaire (GPAQ). The World Health Organization developed the GPAQ to measure physical activity levels across three domains occurring on a typical day [work around the home and in the workforce, transportation (biking and walking to get from place to place) and recreation] in culturally

4 Trost SG, McIver KL, Pate RR. Conducting accelerometer-based activity assessments in field-based research. *Med Sci Sports Exerc.* 2005; 37 (suppl):S531-43.

Table 1 Protocol for Device Follow-Up Accelerometer Data Collection

Accelerometer Data Parameter	Protocol for study
Device	Actigraph GT3X
Monitoring period	7 days, beginning on Thursday and ending on Wednesday
Placement	Hip-mounted
Frequency of recording (epoch)	Every 10 seconds
Non-wear time criteria	60 consecutive minutes of 0 counts, with allowance for 1-2 minutes of counts between 1 and 100
Minimal wear time per day	600+ minutes (or 10 hours)
Minimal number of days worn	4 or more
Outcome variables	Minutes sedentary, minutes of light, moderate or vigorous activity and minutes of moderate or vigorous activity in occurring in 10 minute bouts.
Activity count cut points	Sedentary <100 counts/minute Light = 100 – 2019 counts/minute Moderate = 2020 – 5998 counts/minute Vigorous = 5999+ counts/minute

diverse populations.[5, 6] The GPAQ used for the PAT interviews (Appendix A) was slightly modified to obtain information on activity as part of paid work separately from home work, rather than on all work (including unpaid work, study/training and household chores) combined.

The GPAQ collects information on the frequency and duration of moderate and vigorous activity in the previous seven days.[7] Respondents were asked if, in the past seven days, their work (if employed), house work, travel and recreation time involved physical activity. If so, they were then asked to report the number of days the activity increased breathing or heart rate (considered as moderate physical activity) and the average time spent doing the activity each day. Additionally, in the workforce and recreational domains, respondents were asked to report the number of days and average length of time spent doing vigorous activity, defined as activity that caused a large increase in breathing or heart rate. For each domain, total weekly moderate and vigorous activity were calculated by multiplying the number of days the respondent engaged in physical activity by the daily average moderate or vigorous minutes of physical activity. As vigorous activity requires approximately twice the energy expenditure of moderate activity,[1] all vigorous minutes were multiplied by two to create "moderate-equivalent minutes." Activity across all domains was summed to create a variable of total weekly moderate-equivalent minutes for each participant.

If a participant reported a daily average of 960 minutes (16 or more hours) or more of activity in any domain or if there were any inconsistent values in a participant's response (such as reporting greater than seven days of activity in a week or reporting zero days of activity but then having a value greater than zero for minutes or hours of activity), the participant was not included in the analysis, following the analytic guidelines established by the World Health Organization.[7]

5 Armstrong T, Bull F. Development of the World Health Organization Global Physical Activity Questionnaire (GPAQ). *J Public Health.* 2006; 14: 66–70
6 Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): nine country reliability and validity study. *J Phys Act Health.* 2009; 6(6):790-804.
7 World Health Organization. Global Physical Activity Questionnaire and Analysis Guide. Retrieved from: http://www.who.int/entity/chp/steps/resources/GPAQ_Analysis_Guide.pdf. Date accessed: 4/9/2010.

Measuring physical activity by accelerometer

Participants wore accelerometers for one week throughout the day except when sleeping, swimming or bathing. Since they wore the devices at home and while working, traveling and during recreational time, the data are comparable to the self-report data about activity in these domains collected by the GPAQ.

To process the accelerometer data for analysis, activity thresholds from the Centers for Disease Control and Prevention's National Health and Nutrition Examination Survey were used to assign minutes as moderate or vigorous. These thresholds were established in calibration studies that compared activity counts recorded by accelerometers to measured energy expenditure during treadmill or track walking and running.[8–11] All accelerometer minutes that ranged from 2,020 through 5,998 counts per minute were considered moderate and minutes at or above 5,999 counts per minute were considered vigorous. Like the survey data, vigorous minutes were multiplied by two to create moderate-equivalent minutes. To obtain weekly physical activity values for participants with four to six days of data (n=336), the minutes of moderate-equivalent activity on valid days (≥10 hours of wear time) were summed and divided by the number of valid days of wear time to create a daily average moderate-equivalent activity variable. The daily average was multiplied by seven to create a weekly total.

Comparing measures of physical activity

First, the overall minutes of physical activity from self-reported (across all measured domains) and accelerometer data were compared to determine the extent to which values from the two measures were equal. For the purposes of this analysis, the accelerometer data were considered the gold standard for absolute measurement of minutes of activity. Descriptive statistics of the minutes of physical activity recorded from each measure, including distribution histograms, means and medians, were used to examine the level of agreement between the absolute measures from the accelerometer data and the self-reported data.

Next, the difference between self-reported activity and accelerometer-measured activity was compared across demographic subgroups of the sample to identify any participant characteristics related to errors in reporting. This analysis compared minutes of activity across subgroups to understand whether relative differences in activity levels would yield the same results using self-reported versus accelerometer data.

Each participant's weekly moderate-equivalent minutes of activity by self-report and by accelerometer were first ranked as a percentile (from 1-100%) relative to the values reported by others in the sample. For each participant, the ranked values for self-report and accelerometer were subtracted from one another to quantify how much and in what direction (over or under) the relative percentile from self-report differed from the relative percentile based on accelerometer data. By using the ranked values, the differences were not influenced by outlying values (outliers discussed more in the Results section).

For example, one participant reported 360 weekly moderate-equivalent minutes of activity, which, relative to the others in the sample, is at the 42nd percentile. By accelerometer, the same participant recorded 126 weekly moderate-equivalent minutes of activity, which ranked at the 36th percentile relative to the others in

---

8 Brage S, Wedderkopp N, Franks PW, Andersen LB, Froberg K. Reexamination of validity and reliability of the CSA monitor in walking and running. *Med Sci Sports Exerc.* 2003; 35:1447-54.
9 Freedson PS, Melanson E, Sirard J. Calibration of the Computer. Science and Applications, Inc. accelerometer. *Med Sci Sports Exerc.* 1998; 30:777-81.
10 Leenders NY, Sherman WM, Nagaraja HN, Kien CL. Evaluation of methods to assess physical activity in free-living conditions. *Med Sci Sports Exerc.* 2001; 33:1233-40.
11 Yngve A, Nilsson A, Sjostrom M, Ekelund U. Effect of monitor placement and of activity setting on the MTI accelerometer output. *Med Sci Sports Exerc.* 2003; 35:320-6.

the sample. Consequently, the difference between ranks for this participant was six (or a +6 percentage point difference), suggesting that the participant slightly overreported his/her activity relative to others in the sample. Likewise, the difference between ranks for another participant was –60 (or a –60 percentage point difference), suggesting this participant greatly underreported his/her activity.

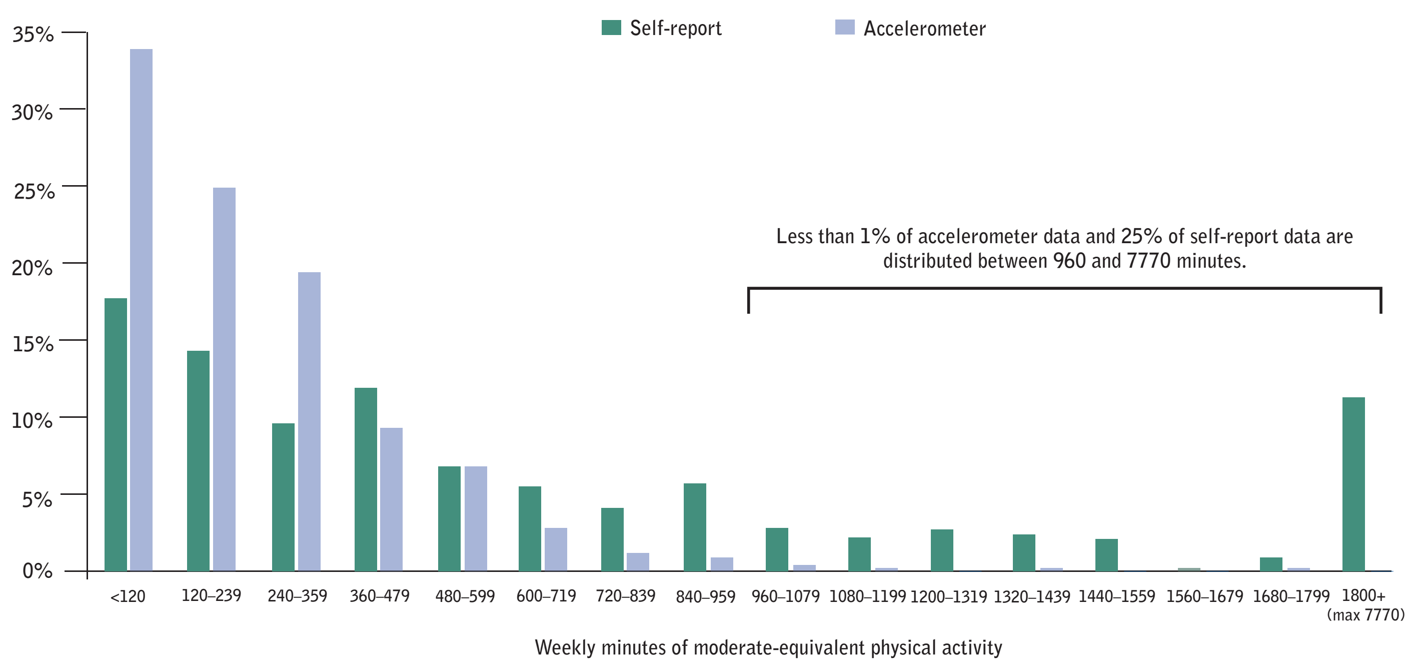
To determine if the error in reporting was statistically related to any participant characteristics, a linear regression model was run with the difference in ranks as the outcome variable. The relationship of this difference in ranks was tested across demographic groups (age, race/ethnicity, gender, education level), geographic groups (borough of residence, population density based on zip code of residence, residence in a high-obesity neighborhood) and groups with health-related conditions (weight status, chronic disease). All variables were included together in one model.

Although a minimum of 10 hours of wear time on at least four days was required to be included in the analytic dataset, the amount of time participants wore the accelerometer each day greatly varied. (The maximum value was an average of 18 hours per day.) Therefore, the total minutes the participants wore the accelerometer device also was included in the model, since participants who wore the device for longer periods had a greater opportunity to accumulate minutes of physical activity.

Results

Figure 1 shows the distribution of weekly moderate-equivalent minutes of activity from self-reported and accelerometer measures. The distributions show that, by both measures, physical activity varied widely in the sample, with a large number of individuals having little activity throughout the week and a smaller percentage having relatively high levels of activity. Self-reported activity had a much wider range of values than accelerometer and contained substantially higher values.

Figure 1 Distribution of weekly moderate-equivalent minutes as measured by self-report and accelerometer

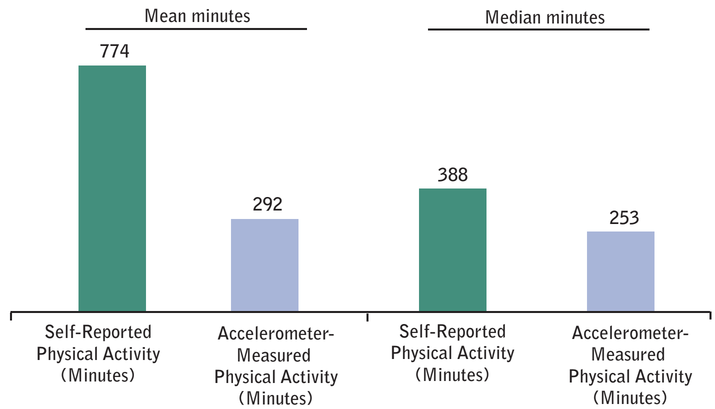


Source: NYC Physical Activity and Transit Survey 2011

For example, based on the self-reported data, 25% of the sample had 960 or more minutes of physical activity (with a maximum of 7,770 minutes) while with accelerometer data less than 1% of adults had this amount of activity (with a maximum of 1,739 minutes). Only three individuals (<1%) had zero minutes of accelerometer-measured activity compared with 41 individuals (6%) reporting zero minutes of activity across all domains of the GPAQ.

If the data were normally distributed, the distribution would look like a bell-shaped curve. Instead, the distribution is asymmetric, or skewed. As a result, analyzing the relationship between the two measures using methods that rely on mean values is misleading. As an alternative, using the ranked data and the median value (the value at the 50th percentile) as an average measure removes the influence of the skewed data. Figure 2 illustrates the influence of outlying values on estimates of mean minutes of activity from self-report. The estimate of mean weekly minutes of activity was substantially higher by self-report than it was by accelerometer measurement. However, median values were much closer, suggesting that outlying values influenced the level of agreement between the two measures. Even using medians, overall physical activity minutes were higher when based on self-report than on accelerometer,* indicating that self-reported physical activity minutes were overestimated.

Figure 2 Physical Activity by Self-Report and Accelerometer — Mean and Median Weekly Minutes



Source: NYC Physical Activity and Transit Survey 2011

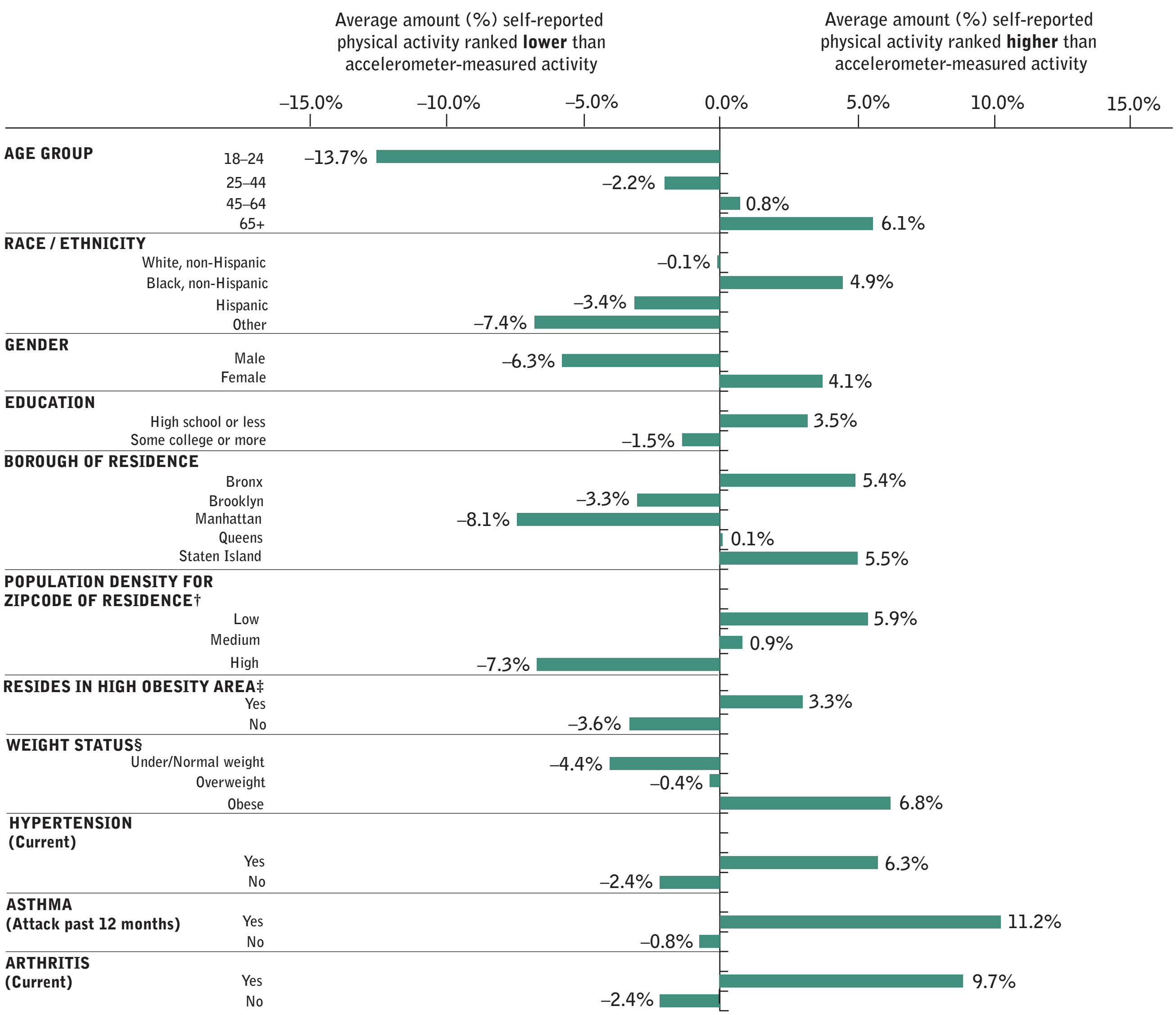
To examine whether overestimation was constant across groups, accelerometer percentile ranks were subtracted from self-reported ranks (Figure 3). Positive values represent overreporting, negative values represent underreporting. The results highlight that over- and underreporting was not consistent across all subgroups. For example, on average, the minutes of self-reported activity for participants in the youngest age group (18 to 24 years) were ranked 13.7 percentage points lower than their minutes of activity by accelerometer measurement, while self-reported minutes of activity from participants in the oldest age group (65 years and older), on average, ranked 6.1 percentage points higher than their accelerometer-measured physical activity.

Results from the regression analysis identified the groups that were statistically different from one another in their over- or underreporting, while holding constant the amount of time the accelerometers were worn. Only age, gender and the amount of time the accelerometer was worn were statistically related to the differences in ranks (Table 2). The parameter estimates in Table 2 represent how far apart the ranks were in their over- and underestimation. For example, the parameter estimate for gender was 11.3. Therefore, if the rank of males' self-report and accelerometer activity were the same (the percentile rank of the self-report data minus the percentile rank of the accelerometer data = 0) then for females the activity rank by self-report would be 11.3 percentage points higher than the rank from accelerometer. Similarly, if females had 0% difference in ranks, males, on average, would under report their rank by 11.3 percentage points. This indicates that the relationship

** Wilcoxon matched-pair signed-rank test were statistically significant (p<.05).*

between self-reported and accelerometer-measured activity was different for males and females. Females' self-reported activity ranked higher than their accelerometer-measured activity while males' accelerometer-measured activity ranked higher than their self-reported activity. This pattern is also illustrated in Figure 4a and 4b. Although the distribution of values by self-report is relatively similar for

Figure 3 Percent difference between ranked self-reported and ranked accelerometer measured physical activity by participant characteristics.*



** Weekly minutes of activity by self-report and by accelerometer were ranked as a percentile (from 1-100%) for each participant. Ranked values from accelerometer were subtracted from ranked values from self-report to quantify how much and in what direction (over or under) the relative percentile from self-report differed from the relative percentile seen in the accelerometer data. The percentage in this figure represents how far apart the ranks are in their over and underestimation for each subgroup. Since accelerometer ranks were subtracted from self-report, positive values represent overreporting, while negative values represent underreporting.*

† Population density: NYC zip codes (ZCTAs) were grouped by population density based on population from the Census 2010 into low density (bottom 25%), medium density (middle 25%-75%) and high density (top 25%) based on ZCTA-level population per square mile.

‡ High obesity areas were defined from multiple years of the Community Health Survey (CHS), and represent continuous United Hospital Fund (UHF) areas with the highest proportion of obesity. There are four separate high obesity areas: Northern Staten Island, East/Central Brooklyn, the South Bronx and Harlem, and Eastern Queens.

§ Weight status is based on Body Mass Index (BMI) calculated from self-reported height and weight from the NYC PAT: underweight is BMI less than 18.5; healthy weight BMI 18.5-24.9; overweight, BMI 25-29.9; obese, BMI 30+.

Source: NYC Physical Activity and Transit Survey 2011

Table 2 **Regression coefficients predicting difference between activity ranks from self-reported and accelerometer measurement*†**

	Parameter Estimate	Standard Error	P
Age			
18 - 24	−18.7	6.7	0.005
25 - 44	−8.2	4.1	0.04
45 - 64	−6	3.7	0.11
65+	ref		
Race/Ethnicity			
White, non-Hispanic	ref		
Black, non-Hispanic	1.8	3.6	0.62
Hispanic	−4.6	3.8	0.22
Gender			
Male	ref		
Female	11.3	2.8	<.0001
Education			
High school or less	ref		
Some college or more	−2.1	3	0.47
Borough of Residence			
Bronx	8.9	5.2	0.09
Brooklyn	1.4	5	0.78
Manhattan	ref		
Queens	1.6	5.2	0.76
Staten Island	6.6	7	0.34
Population Density for Zipcode of Residence‡			
Low	ref		
Medium	−3.7	4.9	0.45
High	−9.3	5.9	0.12
Reside in High Obesity Area§			
Yes	3.7	2.9	0.20
No	ref		
Obesity Status**			
Under/Normal weight	ref		
Overweight	2.3	3.2	0.37
Obese	5.8	3.4	0.08
Hypertension (current)			
Yes	ref		
No	−0.8	3.2	0.81
Asthma (attack in past 12 months)			
Yes	ref		
No	−7.7	5.1	0.13
Arthritis (current)			
Yes	ref		
No	−4.3	3.6	0.23
Time Wore Device			
Total minutes per week	−0.00475	0.001	0.01

* *Analysis done using Ordinary Least Squares linear regression model. All variables included in one model.*

† *Weekly minutes of activity by self-report and by accelerometer were ranked as a percentile (from 1-100%) for each participant. Ranked values from accelerometer were subtracted from ranked values from self-report to quantify how much and in what direction (over or under) the relative percentile from self-report differed from the relative percentile seen in the accelerometer data. The percentage in this figure represents how far apart the ranks are in their over and underestimation for each subgroup. Since accelerometer ranks were subtracted from self-report, positive values represent overreporting, while negative values represent underreporting.*

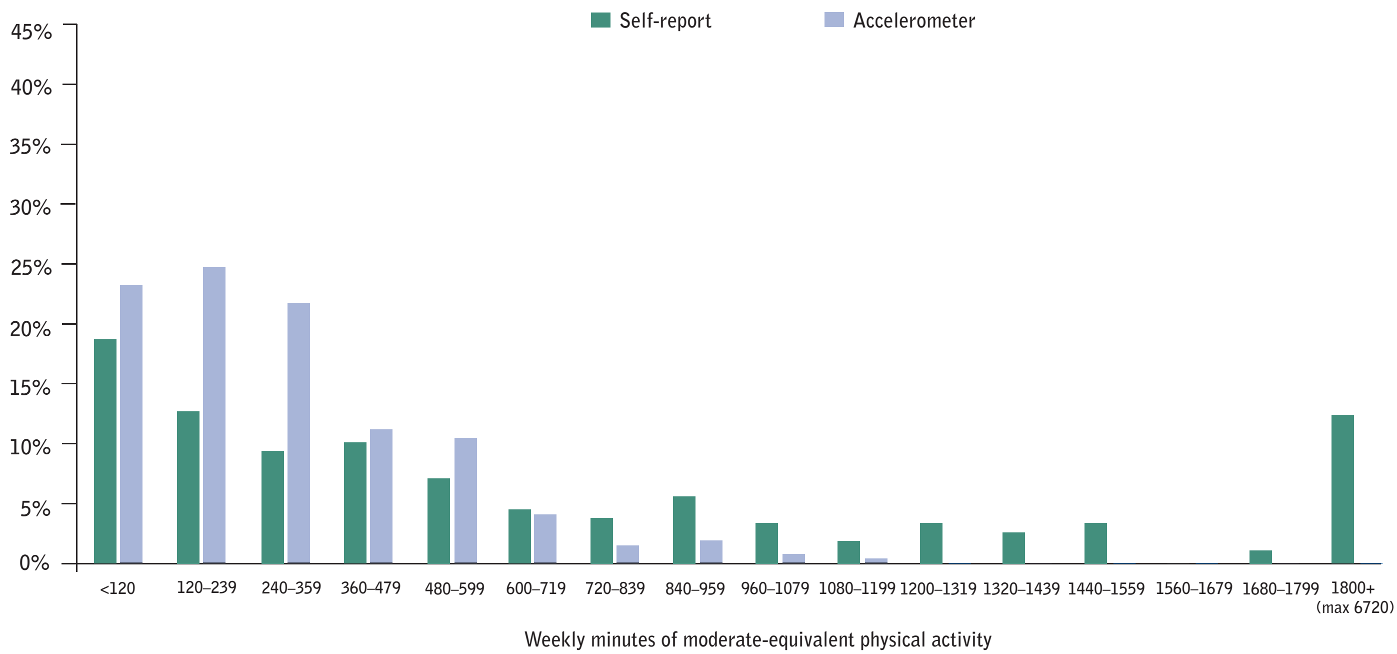
‡ *Population density: NYC zip codes (ZCTAs) were grouped by population density based on population from the Census 2010 into low density (bottom 25%), medium density (middle 25%-75%) and high density (top 25%) based on ZCTA-level population per square mile.*

§ *High obesity areas were defined from multiple years of the Community Health Survey (CHS), and represent continuous United Hospital Fund (UHF) areas with the highest proportion of obesity. There are four separate high obesity areas: Northern Staten Island, East/Central Brooklyn, the South Bronx and Harlem, and Eastern Queens.*

** *Weight status is based on Body Mass Index (BMI) calculated from self-reported height and weight from the NYC PAT: underweight is BMI less than 18.5; healthy weight BMI 18.5-24.9; overweight, BMI 25-29.9; obese, BMI 30+.*

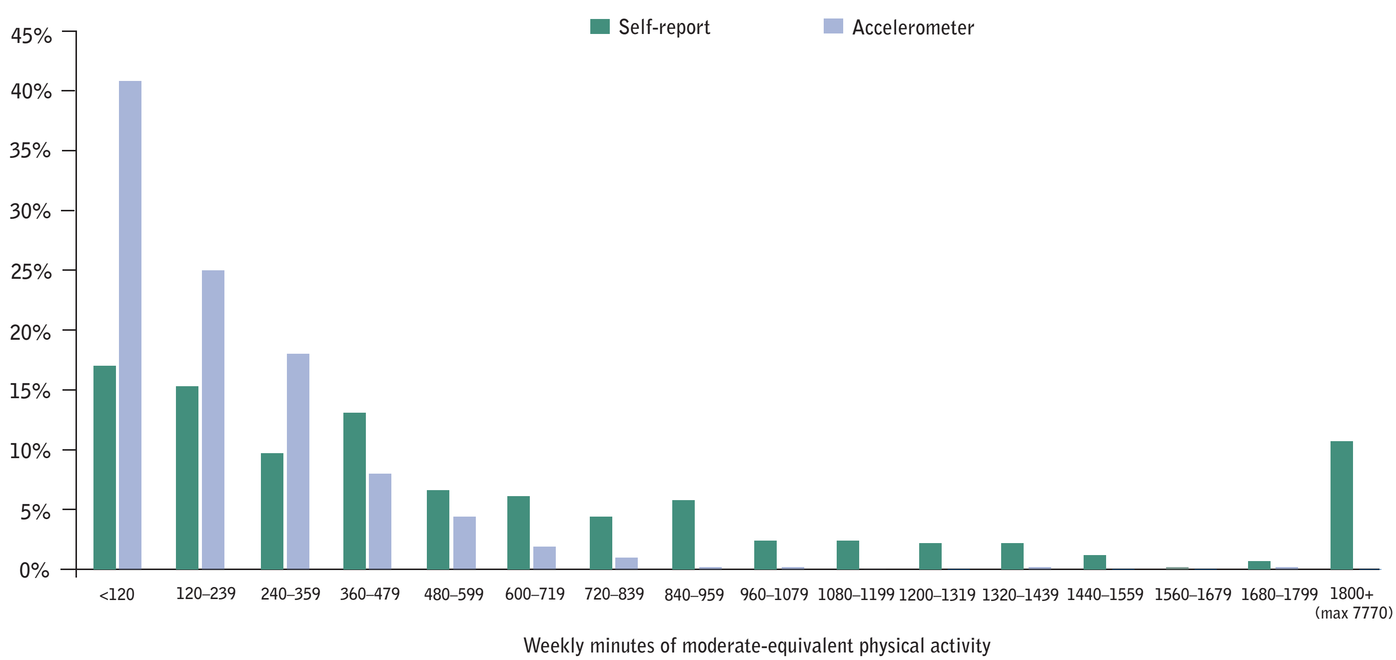
Source: NYC Physical Activity and Transit Survey 2011

Figure 4a **Distribution of weekly moderate-equivalent minutes as measured by self-report and accelerometer for males**



Source: NYC Physical Activity and Transit Survey 2011

Figure 4b **Distribution of weekly moderate-equivalent minutes as measured by self-report and accelerometer for females**



Source: NYC Physical Activity and Transit Survey 2011

males and females, the distribution of accelerometer data is markedly different. Weekly minutes of accelerometer-measured activity clustered at the lower end of the distribution for females, while the distribution for males was more aligned with the self-reported data. Most notable is that the percent of males and females that reported fewer than 120 minutes of weekly activity is nearly identical (18.7% and 17.0% respectively), while only 23% of males and 40% of females had accelerometer-measured activity at this level.

Recommendations

The PAT data provide a snapshot of physical activity that is both robust enough to make population-level estimates of physical activity across various domains in a very large, urban environment and detailed enough to determine how accurate each individual's measure of activity is, down to the minute. The data from both self-reported and objective measurement contribute unique information that can advance public health understanding of the factors that influence physical activity and health outcomes. Based on our comparison of self-reported to accelerometer data on absolute levels of physical activity, several recommendations should be considered when analyzing self-reported data.

Recommendation 1: Use median minutes of physical activity when reporting estimates of average activity levels

Physical activity data were highly skewed, regardless of method of collection, and require the use of medians as an average measure.* This reporting method is a common way to address skewed physical activity data often seen in national or large population datasets.[12] Also, in reporting on the development of the GPAQ, Armstrong and Bull have suggested medians as a continuous outcome measure.[4] Moreover, though medians were generally higher in self-reported data than in accelerometer, the difference between median self-reported physical activity minutes and accelerometer minutes was not as extreme as the difference in mean minutes.

Recommendation 2: Stratify estimates by gender and, if possible, age

Females were ranked higher by their self-report measure than by their accelerometer measure; the opposite was true for males. A similar pattern was also observed across age groups. Older participants were ranked higher by self-report than by accelerometer and the opposite was true for younger participants. This suggests that researchers using self-reported physical activity data should understand that older age groups and women may overreport their physical activity more than younger people and men.

Given that self-reported data and accelerometer data have different patterns of agreement across groups, stratifying self-report estimates by gender and/or age when making inferences about relative relationships between physical activity and other factors will restrict the bias to each group. For example, although physical activity values for males may be underreported, if the sample is stratified by gender, this error is held constant within each group (i.e., males and females) and observed associations between physical activity and other factors are more likely to be valid. Likewise, it should be noted that, assuming the error remains the same across time, the error would be held constant when comparing values over time even without stratification.

Recommendation 3: Self-reported physical activity data can be a cost-effective surveillance tool, capturing activity domains that accelerometers cannot

The analyses summarized in this report show inconsistencies in estimates of absolute physical activity levels of adult New Yorkers as calculated from self-report

* *No additional steps were taken to remove outliers beyond what is advised in the GPAQ analytic guidance document, as this is not the convention and would further reduce the sample size. Although medians were chosen as the best method of presentation for physical activity estimates for this analysis, there are other methods available for analyzing non-normally distributed data.*

12 Karolinska Institutet. Guidelines for the data processing and analysis of the International Physical Activity Questionnaire. Retrieved from: http://www.ipaq.ki.se/scoring.pdf. Date accessed: 5/10/2010.

versus accelerometer data. Recommendations on how to better align the estimates of minutes of physical activity between each measure are provided above. However, to assess the utility of self-reported physical activity data, researchers and evaluators should consider more than just the alignment of survey data with accelerometer data.

Accelerometers objectively measure individuals' movement and capture data on physical activity from points in time that would be difficult for survey respondents to recall accurately. Yet, measuring physical activity by accelerometer can be cost-prohibitive, making self-report a more practical option for routine public health surveillance. Moreover, self-reported physical activity can give details on the type, context and purpose of physical activity, which may be more relevant to certain surveillance, research and evaluation needs than only gross recorded minutes of activity without such context.[13] For example, the GPAQ questions regarding recreational activity or the questions on leisure-time activity in the Health Department's Community Health Survey[14] can be used to specifically track discretionary physical activity. Questions regarding walking and biking can be used to track physical activity related to active transportation.

As the definition of physical activity can range from planned exercise to overall active living, information on self-reported physical activity is essential to understanding the *behavior* of physical activity, and the assessment of its validity should not be limited to its association with accelerometer-measured *movement*.

Limitations and Strengths

The recommendations included in this report will improve the accuracy of self-report estimates. There are, however, some limitations to the analyses. The GPAQ asked about activity that occurs in bouts lasting at least 10 minutes, as these bouts of self-reported activity are the basis for national physical activity guidelines.[1] Studies on the association between the optimal lengths of accelerometer-measured activity for improved health are not available. Thus, accelerometer activity was not restricted to bouts, which is consistent with accelerometer validations of the GPAQ reported elsewhere.[6] Also, the GPAQ asked about the previous week, while accelerometers were worn after the survey was administered. The average time between survey completion and start of accelerometer participation was three weeks, making it unlikely that differences were due to seasonal effects. Bout restrictions and different time periods may have introduced some error in comparing the measures of physical activity.

Weighting vigorous minutes of physical activity to reflect greater energy expenditure than moderate minutes is necessary to categorize individuals as "sufficiently active" by national recommendations[1] and by GPAQ analytic guidelines.[7] However, a larger gap between self-report and objective measurement will be seen if vigorous minutes are disproportionately overreported. In the PAT sample of moderate and vigorous activity, there was a greater proportion of vigorous minutes in self-reported data (16% vigorous, 84% moderate) than in accelerometer data (4% vigorous, 96% moderate). Therefore, the differences seen between measured minutes may be influenced more by overreporting of vigorous activity, rather than overall overreporting. Future studies should assess the effects of weighting vigorous activity minutes.

In addition, accelerometers worn on the hip do not move during stationary activity (e.g., weight lifting) or biking and cannot be worn while people are swimming. Although these activities made up a small portion of the accelerometer data relative to the rest of the movement that occurs throughout the week in this study, self-reported time spent engaging in them may not have been captured by the accelerometers.

Even with these limitations, the methods of data collection used for the PAT are among the best available, representing state-of-the-art science in physical activity assessment. The data can be used to generate population estimates of both self-reported and objectively-measured physical activity in a culturally diverse urban population, which is rarely available in public health surveillance, evaluation or research. The survey data additionally provide a wealth of in-depth information on physical activity behaviors, built environments and active transportation, which is usually limited to small modules in surveillance surveys.

For this reason, the data should be used to thoroughly examine physical activity, while acknowledging the caveats discussed here when reporting analyses.

13 Troiano RP, Pettee Gabriel KK, Welk GJ, Owen N, Sternfeld B. Reported physical activity and sedentary behavior: why do you ask? *J Phys Act Health.* 2012; 9(Suppl 1):S68-75.

14 For more information on the NYC Community Health Survey, visit nyc.gov/health/survey.

Recommendation 4: If physical activity intensity is the focus of analysis, consider the differential experience of activity across groups

Self-reported survey data reflect respondents' perceptions of their own physical activity behaviors relative to others. The difference in reporting seen in this analysis is likely related in part to the objective and subjective natures of the accelerometer and the GPAQ, respectively. The GPAQ asks participants for their subjective perceptions of activity, relative to their own physiological signs (e.g., increased breathing and heart rate). An activity that increases one's heart rate is likely different across certain groups (e.g. younger and older adults, normal weight and obese, those with or without a chronic health condition, etc.). Yet, by definition, objective measures of activity are not categorized by perceptions of intensity. Instead, an absolute threshold of movement, which is related to a metric of energy expenditure, must be crossed to be considered moderate or vigorous activity. In other words, although it may be more difficult for an older individual to walk at a brisk pace of 3.0 miles per hour than a younger individual, only by doing so are either "moderately active" by an objective measure. The different conceptualization of activity intensity in the GPAQ measure and the accelerometer measure may result in a gap between the values captured by each. When deciding what to track in public health surveillance or what to monitor in intervention evaluation and research, it is important to recognize that each method may have a different measurement characteristic, which needs to be carefully defined when operationalizing the concept of "physical activity".[13]

Conclusion

Measurement of physical activity on a population level has proven to be a challenging task in public health research and surveillance, with a variety of factors contributing to overestimation of absolute levels of activity by self-report.[15] Ongoing research, like this study, is working toward better methods for collecting self-reported data on physical activity, informed by comparison with accelerometer data and across varying activity scales like the GPAQ.[16] Although self-reported data on minutes of activity do not accurately quantify absolute amounts of physical activity, they can be used to evaluate relative differences in activity levels over time and among subgroups with context of domain and purpose. Such data are essential to the development and evaluation of public health interventions to increase physical activity. The Health Department will continue to conduct physical activity surveillance using self-reported data and, when available, accelerometer data, employing the recommendations above. We will also continue to evaluate varying survey collection methods to maximize accuracy and best inform public health programs and policy.

15 Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. *Br J Sports Med.* 2003; 37(3):197–206.
16 Bowles HR. Measurement of active and sedentary behaviors: closing the gaps in self-report methods. *J Phys Act Health.* 2012; 9 (Special issue).

NYC Health
New York City Department of Health and Mental Hygiene
42-09 28th Street, CN6, Queens, NY 11101

April 2013

Mayor: Michael R. Bloomberg
Commissioner of Health and Mental Hygiene: Thomas A. Farley, MD, MPH
Division of Epidemiology: Carolyn Greene, MD, Deputy Commissioner

Suggested Citation: *Wyker B, Bartley K, Holder-Hayes E, Immerwahr S, Eisenhower D, Harris TG. Self-Reported and Accelerometer-Measured Physical Activity: A Comparison in New York City. New York City Department of Health and Mental Hygiene: Epi Research Report, April 2013; 1-12.*

Acknowledgements: *Thanks to the following individuals who contributed to the conceptualization and review of this product: Tamar Adjoian, Kevin Konty, Karen Lee, James McClain, Jennifer Norton, Carolyn Olson, Jim Sallis, Richard Troiano, Stewart Trost, Gretchen Van Wye.*

Epi Research Report

New York City Department of Health and Mental Hygiene April 2013

APPENDIX A
Modified Global Physical Activity Questionnaire (used in PAT)

The Physical Activity and Transit (PAT) survey, conducted by the New York City Department of Health and Mental Hygiene, used a version of the Global Physical Activity Questionnaire (GPAQ) to measure self-reported physical activity (PA). The GPAQ was modified slightly for telephone administration, and activity related to paid work and housework was asked separately, as is the case with the long version of the International Physical Activity Questionnaire (IPAQ).

DK = Don't know

Next I am going to ask you about the time you spent in the last 7 days doing different types of physical activity (if employed "when you were at work,") when you were at home, engaged in recreation, and going from place to place.

I will ask you about physical activities that caused increases in your breathing or heart rate and that lasted for at least 10 minutes at a time. Do not include light levels of activity that did not increase your breathing or heart rate.

LABOR FORCE PA [If employed]

1 Think first about the time you spent working in the last 7 days.

Did your work involve any physical activity that caused an increase in your breathing or heart rate such as from carrying light loads, lifting heavy loads, or digging or construction work?

[If no, DK or refused, skip to Home PA]

2 During the last 7 days, on how many days did your work activities cause increases in your breathing or heart rate?

Number of days _____ (range 1-7)

[If DK or refused, skip to Home PA]

3 On an average day during the last 7 days, how many hours or minutes did you spend at work doing those activities that increased your breathing or heart rate?

[PROBE IF DIFFICULTY: Think about the last day at work when you did activities that made your heart rate or breathing increase.]

Number of minutes _____ (range 0-59) or hours______ (range 0-24) [Skip to 4]

[If DK, 3a]

[If refused, skip to Home PA]

3a Would you say that it was less than 30 minutes, 30 to 60 minutes, or more than 60 minutes?

[Skip to Home PA]

4 How many, if any, of those (fill hours or minutes from above) were vigorous activity that caused a large increase in your breathing or heart rate?

Number of minutes _____ (range 0-59) or hours______ (range 0-24)

HOME PA

Now think about the physical activities you have done in the last 7 days in and around your home, like housework, yard work, and caring for your family. Exclude light activities that do not increase your breathing and heart rate.

5. In the last 7 days, did you do work around the house for at least 10 minutes at a time that caused an increase in your breathing or heart rate? Some examples would be yard work, sweeping, washing windows, or lifting children.

[If no, DK or refused, skip to Travel PA]

6. During the last 7 days, how many days did you do activities around the home that caused an increase in your breathing or heart rate?

Number of days _____ (range 1-7)

[If DK or refused, skip to Travel PA]

7. On an average day during the last 7 days, how many hours or minutes did you spend doing these activities

Number of minutes _____ (range 0-59) or hours______ (range 0-24) [Skip to Travel PA]

[If DK, 7a]

[If refused, skip to Travel PA]

7a Would you say that it was less than 30 minutes, 30 to 60 minutes, or more than 60 minutes?

TRAVEL PA

Now I would like to ask you about the usual way you travel to and from places such as work, shopping, or school.

[WALKING PA]

8 During the last 7 days, did you walk for at least 10 minutes at a time to get to and from places such as work, shopping, or other activities?

[If no, DK or refused, skip to Bike PA]

9 During the last 7 days, on how many days did you walk to get to and from places?

Number of days _____ (range 1-7)

[If DK or refused, skip to Bike PA]

10 How many hours or minutes did you spend walking to get from place to place on an average day during the last 7 days?

[READ IF NEEDED: Think about yesterday if an average day is too difficult to determine.]

Number of minutes _____ (range 0-59) or hours______ (range 0-24) [Skip to Bike PA]

[If DK, 10a] [If refused, skip to Bike PA]

10a Would you say that it was less than 30 minutes, 30 to 60 minutes, or more than 60 minutes?

[BIKE PA]

11 During the last 7 days, did you use a bicycle to get to and from places?

[If no, DK or refused, skip to Recreation PA]

12 During the last 7 days, on how many days did you bicycle to get to and from places?

Number of days _____ (range 1-7)

[If DK or refused, skip to Recreation PA]

13 How many hours or minutes did you spend bicycling to get from place to place on an average day during the last 7 days?

[PROBE IF DIFFICULTY: Think about the last day when you used your bicycle to go from one place to another.]

Number of minutes _____ (range 0-59) or hours______ (range 0-24) [Skip to Recreation PA]

[If DK, 13a]

[If refused, skip to Recreation PA]

13a Would you say that it was less than 30 minutes, 30 to 60 minutes, or more than 60 minutes?

RECREATION PA

Now I would like to ask you about sports, fitness and recreational activities. Exclude the activities you have told me about and focus on sports, fitness and recreation.

14 During the last 7 days, did you do any sports, fitness, or recreational activities that caused an increase in your breathing or heart rate? Examples would be swimming, working out, or jogging.

[If no, DK or refused, skip to Sedentary]

15 During the last 7 days, on how many days did you do sports, fitness or recreational activities that increased your breathing or heart rate?

Number of days _____ (range 1-7)

[If DK or refused, skip to Sedentary]

16 On an average day during the last 7 days, how many hours or minutes did you spend doing these sports, fitness or recreational activities?

Number of minutes _____ (range 0-59) or hours______ (range 0-24) [Skip to 17]

[If DK, 16a]

[If refused, skip to Sedentary]

16a Would you say that it was less than 30 minutes, 30 to 60 minutes, or more than 60 minutes?

[Skip to Sedentary]

17 How many, if any, of those (fill hours or minutes from above) doing sports, fitness, or recreational activities were vigorous activity that caused a large increase in your breathing or heart rate?

Number of minutes _____ (range 0-59) or hours______ (range 0-24)