Validity of two brief physical activity questionnaires with accelerometers among African-American women

Rodney P. Joseph[1] [Postdoctoral Fellow], **Colleen Keller**[2] [Professor], **Marc A. Adams**[3] [Assistant Professor], and **Barbara E. Ainsworth**[4] [Professor]

[1]College of Nursing and Health Innovation, Arizona State University, Phoenix, Arizona, USA

[2]College of Nursing and Health Innovation, Arizona State University, Phoenix, Arizona, USA

[3]School of Nutrition and Health Promotion, Arizona State University, Phoenix, Arizona, USA

[4]School of Nutrition and Health Promotion, Arizona State University, Phoenix, Arizona, USA

Abstract

Aim—To evaluate the validity of the Stanford Brief Activity Survey (SBAS) and Exercise Vital Sign (EVS) questionnaire against accelerometer-determined time in moderate-to-vigorous physical activity (MVPA) among African-American (AA) women.

Background—Limited research has evaluated the validity of brief physical activity (PA) questionnaires among AA women. Since the validity of PA questionnaires may differ among members of varying racial/ethnic groups, research is needed to explore the validity of self-report PA measures among AA women.

Methods—A total of 30 AA women [M age = 35.5 ± 5.3; M body mass index (BMI) = 31.1 ± 7.8] wore ActiGraph GT3X + accelerometers (ActiGraph, LLC, Pensacola FL, USA) for seven days and completed both the SBAS and EVS at two different assessment periods (T1 and T2). Criterion validity was calculated using Spearman's rank order correlations between each questionnaire score and accelerometer-measured MVPA. Sensitivity, specificity, and positive and negative predictive values were calculated using accelerometer-measured MVPA as the criterion to determine the ability of each questionnaire to predict whether or not a participant was meeting the 2008 US PA Guidelines

Findings—Spearman correlation coefficients between questionnaire scores and minutes of accelerometer-measured MVPA were low (EVS, $r = 0.27$ at T1 and $r = 0.26$ at T2; SBAS, $r = 0.10$ at T1 and $r = 0.28$ at T2) and not statistically significant (Ps > 0.05). The EVS had sensitivity, specificity, and negative and positive predictive values of 27, 89, 59, and 68% at T1 and 33, 74, 38, and 70% at T2, respectively. The SBAS had sensitivity, specificity, and negative and positive

predictive values were 18, 79, 33, and 62% at *T*1 and 67, 58, 43, and 79% at *T*2. While both questionnaires may be useful in identifying AA women who do not meet the 2008 PA Guidelines, using the questionnaires to identify AA women meeting the PA Guidelines should be done with caution.

Keywords

black; exercise; female; minority health; physical activity assessment

Introduction

The benefits of physical activity (PA) are well established for the prevention and treatment of various health conditions, including overweight and obesity, cardiovascular disease, type II diabetes, and select cancers (ie, breast and colorectal) (US Department of Health and Human Services, 2008). Current US Guidelines recommend that Americans engage in a minimum of 150 min of moderate-intensity aerobic PA each week in order to receive health benefits and reduce risk for chronic disease (US Department of Health and Human Services, 2008). Despite these recommendations, many Americans, especially minorities and women, are insufficiently active. National data show that ~36% of African-American (AA) women achieve the national PA Guidelines [Centers for Disease Control and Prevention (CDC), 2010]; which is substantially lower than the national average (48%) and to the PA levels White women (46%) (CDC, 2010). AA women are also disproportionally burdened by disease conditions associated with insufficient PA, including overweight and obesity (Ogden *et al.*, 2014), type II diabetes (Kahn *et al.*, 2006; Goossens, 2008; Bluher, 2009; Glauber and Karnieli, 2013), metabolic syndrome (Ervin, 2009), and cardiovascular disease (Mosca *et al.*, 1997; Mosca, 2007; Ashraf and Baweja, 2013; Bastien *et al.*, 2014). The low PA levels and associated disparate chronic disease burden among AA women indicate the need for PA promotion efforts in this high-risk population.

An important aspect for researchers, clinicians, and public health practitioners to consider when evaluating and promoting PA among AA women is the validity of measure used to assess PA. Accurate assessment of PA is critical to identify individuals who could benefit from increasing their PA, as well as to evaluate the effectiveness of PA promotion efforts. Objective PA measures, such as accelerometers and doubly labeled water, are often viewed as the 'gold standard' of PA assessment (Plasqui and Westerterp, 2007; Westerterp, 2009; Plasqui *et al.*, 2013). However, objective PA measures are not always feasible or appropriate for use due to cost and time restraints, as well as the high level participant burden. For example, PA assessment using accelerometers requires participants to wear an activity monitor (usually on their hip or wrist) for an extended period of time (ie, four to seven days) to obtain a valid assessment (Trost *et al.*, 2005). This can be troublesome for participants because they have to remember to wear the accelerometer on a daily basis and be willing to wear it even when they perceive it as intrusive or uncomfortable. The cost of accelerometers is also high, with prices ranging from $250 to $500 for a single device. PA assessed using doubly labeled water has similar barriers, as this method requires participants to drink oxygen and hydrogen enriched water and provide frequent urine, saliva, and/or blood samples during the assessment period (usually between 7 and 21 days; Levine, 2005). The

cost is also substantial, with a typical assessment costing around $1500/participant. Questionnaires, on the other hand, represent a quick, low cost, and low participant burden alternative to assessing PA.

Most PA questionnaires used in the United States have been predominately validated among White populations with little attention paid to the validity of the measures among racial/ ethnic minorities (Sallis and Saelens, 2000; Wolin *et al.*, 2010). Evidence suggests that AA women have different PA preferences than White women [ie, prefer group-based PAs that allow socialization and interaction (eg, group-based walking, dance, aerobics) and have limited interests in performing more individualistic or racquet-based sporting activities (eg, swimming, golf, tennis); Nies *et al.*, 1999; Wood, 2002; Pekmezi *et al.*, 2013] and engage in more of occupational PA (Evenson *et al.*, 2003; He and Baker, 2005). Accordingly, it is possible that the sample activities and recall prompts included in most PA questionnaires do not coincide with the PA preferences or activities performed by AA women. Similarly, most PA questionnaires require high levels of literacy and/or numeracy to understand the questionnaire items and recall activities performed, which may also effect the validity of the measures among AA women since AA women, on average, have lower education levels than their White counterparts (United States Census Bureau, 2014). This suggests that the validity of self-report PA instruments may differ among AA women when compared with predominately White populations.

In a 2010 review of self-report PA measures used to assess PA among US Blacks, Wolin *et al.* (2010) identified only 10 PA questionnaires that have been validated in Black or AA populations. Of these 10 questionnaires, none were brief PA assessment instruments (ie, two items). The shortest PA questionnaires included the Nurse's Health Study (eight-items) (Young *et al.*, 2001) and walking items of the Paffenbarger PA Questionnaire (eight-items) (CDC, 2007). We reviewed the literature to identify additional validation studies of brief PA questionnaires among AA women since the publication of Wolin *et al.*'s (2010) review. Results of our review had similar findings, as we failed to identify any studies validating brief PA questionnaires in this population. The lack of validated brief PA assessment instruments among African American is concerning because brief PA assessment questionnaires have practical applications in research, clinical, and practice settings. For example, in clinical settings, brief PA questionnaires can be used by physicians to quickly identify individuals who need consultation to increase their PA. In research settings, brief PA questionnaires can be used as screening instruments to identify individuals with low PA levels for enrollment in PA promotion interventions. Similarly, epidemiologist can use them to evaluate PA behaviors of large populations.

The purpose of the current study was to evaluate the validity of two brief, two-item PA questionnaires, The Exercise 'Vital Sign' (EVS) (Coleman *et al.*, 2012) and the Stanford Brief Activity Survey (SBAS) (Taylor-Piliae *et al.*, 2006), against accelerometer-measured PA among a sample of AA women residing in the Southwestern United States. Specifically, we assess the ability of each questionnaire to accurately classify moderate-to-vigorous PA status and to identify AA women meeting the 2008 US PA Guidelines of 150 min/week of at least moderate-intensity PA.

Methods

Participants

Participants were 30 AA women enrolled in an eight-week randomized pilot trial evaluating the effects a multi-component, culturally relevant intervention delivered using Facebook and text-messaging to promote PA among AA women versus a non-culturally tailored print-based PA intervention (Joseph *et al.*, 2015). Eligibility for the original pilot study included: (a) self-identified as AA and female, (b) insufficiently active [ie, <150 min/week of moderate-intensity PA as assessed by the short version of the International Physical Activity Questionnaire (IPAQ)], (c) aged 24–49 years, (d) had an active Facebook account, and (e) could to read and write in English. Exclusion criteria included: (a) concurrent participation in another PA promotion program (research or commercial), (b) pregnancy or planned to become pregnant in the next six months, (c) a contraindication to exercise as indicated by the Physical Activity Readiness Questionnaire (Thomas *et al.*, 1992) unless written permission was provided for the participant's primary care physician.

Study design and procedures

Study procedures were approved by the Institutional Review Board of Arizona State University. Participants completed a baseline study visit and an end of study follow-up visit approximately eight-weeks apart. At the baseline study visit (*T*1), informed consent was obtained and demographic, anthropometric, and PA questionnaire data were collected. Participants were then provided an accelerometer to wear during all waking hours for the next seven days. At the end of the seven-day wear period, participants returned their accelerometer to study staff and received $10 for participation. Week eight data collection (*T*2) procedures were similar to baseline procedures. Participants were recruited in February 2014 from the metropolitan area of Phoenix, Arizona, USA and the last participant completed the intervention in June 2014.

Measures

Demographics—Age, income, occupation, and number of children living in the household were collected by brief questionnaire.

BMI—A senior staff member measured participant's height and weight. To ensure consistency, the same staff member conducted measurements for all study participants. Weights were measured to the nearest kilogram using a Tanita TBF-300 A digital scale (Tanita Corporation of American Inc., Arlington Heights IL, USA). Height was measured to the nearest centimeter using a Seca 213 portable stadiometer (Seca North America, Chino CA, USA). BMI was computed as weight in kilograms divided by height in meters squared.

Accelerometer-measured PA—The ActiGraph GT3X + accelerometer (Acti-Graph Corp, LLC, Pensacola FL, USA) was the criterion PA measure for the study. The GT3X + assesses time spent in sedentary, light, moderate, and vigorous activity PA, as well as provides an estimate of daily step counts. Participants were instructed to wear the accelerometer on their right hip during waking hours for a consecutive seven-day period at both baseline and eight-weeks. Accelerometer data were downloaded and processed with

ActiLife Software (version 6.10.0) using a 60-s epoch without the low frequency extension. To be considered as a valid assessment (and subsequently included in data analysis), participants were required to wear the accelerometer for at least 10 h/day on at least four days during the seven-day wear period. Non-wear time was defined as 60 min of consecutive zero counts with allowance of a 2 min spike tolerance of <100 counts/min. Cut-points developed from controlled laboratory experiments were used to estimate minutes of activity performed at various intensity levels: sedentary (0–99 counts/min) (Matthews *et al.*, 2008), light (100–759 counts/min), moderate-lifestyle (760–1951 counts/min) (Matthew, 2005), moderate (1952–5725 counts/min), and vigorous (>5725 counts/min) (Freedson *et al.*, 1998).

SBAS—The SBAS (Taylor-Piliae *et al.*, 2006) is a two-item, self-administered PA questionnaire that can usually be completed in <5 min. The SBAS assesses both occupational (employment activity such as waitressing) and leisure-time PA (such as walking, tennis, or jogging) and classifies respondents' overall PA intensity on a five-point scale as follows: inactive (1), light (2), moderate (3), hard (4), and very hard (5). The SBAS was originally validated among predominately White, English speaking older adults (aged 60–69) (Taylor-Piliae *et al.*, 2006) and has since been validated in other populations, including middle-aged men and women with early onset coronary artery disease (Taylor-Piliae *et al.*, 2007) and post-partum Latinas (Joseph *et al.*, 2014). However, no studies have evaluated the validity of the SBAS among AA women.

EVS—The EVS was developed by Kaiser Permanente at Southern California in response to the American College of Sports Medicine 'Exercise is Medicine' initiative (Coleman *et al.*, 2012). The EVS was designed for clinicians and health care providers to quickly identify patients who were not meeting the national PA recommendations to assist clinicians in promoting PA (Coleman *et al.*, 2012). The EVS questions were incorporated into traditional vital sign measurement (eg, blood pressure, pulse, temperature) by the Kaiser Permanente Health Care System in October 2009 and has been used to assess and promote PA by Kaiser health care providers ever since. The EVS ask participants to report the frequency (days per week) and duration (minutes per day) of moderate-to-vigorous intensity PA performed (eg, like a brisk walk) during the past week. The questionnaire is scored by multiplying the days × minutes of PA performed to create an estimate of minutes per week of at least moderate-to-vigorous PA. The EVS was originally validated for accuracy against national population-based surveillance surveys (ie, NHANES and the California BRFSS) (Coleman *et al.*, 2012). However, similar to the SBAS, no studies have explicitly evaluated the validity of the EVS against accelerometer-measured PA among AA women.

Statistical analysis

Descriptive statistics (ie, means, medians, and percents) were calculated for all demographic, anthropometric, and PA outcomes. T-test were used to determine if there were differences in accelerometer-measured moderate-to-vigorous physical activity (MVPA) among women classified as meeting or not meeting the 2008 US PA Guidelines. Criterion validity of both questionnaires was calculated using Spearman's rank order correlations between each questionnaire score and accelerometer-measured MVPA. Sensitivity,

specificity, and positive and negative predictive values were calculated using accelerometer-measured MVPA as the criterion to determine the ability of each questionnaire to predict whether or not a participant was meeting the 2008 US PA Guidelines. Brief survey items were dichotomized to classify participants as meeting the 2008 US PA Guidelines using the following criteria: (a) EVS indicated 150 min/week of at least moderate-intensity PA, (b) SBAS score was moderate-intensity or greater (ie, moderate, hard, or very hard). In preliminary analyses, we examined whether validity outcomes differed according to group assignment (ie, culturally relevant Facebook and text message intervention versus the non-culturally tailored intervention). Validity outcomes for both intervention groups were similar at both assessment periods; therefore, the groups were combined and results are presented for the whole study cohort as opposed to by intervention group. All analyses were conducted using IBM Statistical Program for the Social Sciences (SPSS) version 21.1 (SPSS Inc., Chicago, IL, USA). Statistical significance was set a P 0.05.

Results

Sample characteristics

Participants ($n = 30$) had a mean age of 35.5 years (SD = 5.3) and most were overweight ($n = 8$; 26.7%) or obese ($n = 14$; 46.7%) (M BMI = 31.1 ± 7.8). The majority ($n = 27$, 90%) had a bachelor's degree or higher and were not married ($n = 19$, 63.3%). Approximately half ($n = 16$, 53.3%) had at least one child living at home. Annual household income was fairly diverse among participants: 48% ($n = 14$) reported making < $50 000, 27% ($n = 8$) reported an income between $50 000 and $100 000, and 27% ($n = 8$) had an income of > $100 000.

Participants' PA patterns

Participants' mean accelerometer wear time was 790.1 min/day (13.1 h/day) at $T1$ and 840.2 min/day (14.0 h/day) at $T2$. Overall, participants engaged in a mean of 128.0 ± 88.4 min/week of MVPA at $T1$ and 127.8 ± 102.9 min/week of MVPA at $T2$ according to accelerometers. Accelerometer-measured minutes/week of MVPA among participants classified as meeting or not meeting the 2008 PA Guidelines by PA assessment measure are presented in Table 1. A total of 11 women (36.7%) were classified as meeting the 2008 US PA Guidelines of 150 min/week of MVPA according to accelerometers at $T1$ and nine (32.1%) as meeting the 2008 US PA Guidelines at $T2$.

According to the EVS, participants reported engaging in 95.3 ± 102.1 min/week of MVPA at $T1$ and 129.7 ± 133.1 min/week at $T2$. Fewer participants reported engaging in 150 min/week of MVPA on the EVS (17% at $T1$ and 14% at $T2$) compared with accelerometer-measured MVPA (Table 1). Participants achieving the national PA Guidelines based on the EVS performed 189 ± 106.6 min/week of accelerometer-measured MVPA at $T1$ and 193.8 ± 87.5 min/week at $T2$. Likewise, those classified as not achieving the national PA Guidelines performed 115.5 ± 81.4 min/week of accelerometer-measured MVPA at $T1$ and 116.8 ± 102.7 min/week of MVPA at $T2$.

Similar to findings of the EVS, fewer women were classified as meeting the US PA Guidelines according to the SBAS in comparison with accelerometer-measured MVPA (20%

at *T*1 and 50% at *T*2). Mean accelerometer-measured MVPA (minutes/week) for participants classified to each SBAS category at both assessment periods are shown in Figure 1. According the SBAS, all participants at the *T*1 assessment were classified into the inactive (*n* = 1), light (*n* = 23), or moderate-intensity (*n* = 6) PA categories (Figure 1a). No participants were classified into the higher intensity PA categories of hard or very hard. At *T*2, participants self-reported engaging in more PA according to the SBAS as all five PA intensity categories of the SBAS were represented by participants (Figure 1b).

Survey validation

Table 2 illustrates the validity, sensitivity, specificity, and positive and negative predictive values for each questionnaire to predict if a participant met the 2008 PA Guidelines based on accelerometer-measured MVPA.

EVS—Spearman's rank correlation coefficients between the EVS and accelerometer-measured MVPA were low and non-significant at both assessment periods (*T*1 $r = 0.27$, $P = 0.15$; *T*2 $r = 0.26$, $P = 0.17$). The EVS had sensitivity, specificity, and negative and positive predictive values of 27, 89, 59, and 68% at *T*1 and 33, 74, 38, and 70% at *T*2, respectively (see Table 2). Participants classified as meeting the 2008 PA Guidelines based on the EVS engaged in 73.2 more min/week of accelerometer-measured MVPA than participants not meeting the guidelines at *T*1 and 77 more min/week of accelerometer-measured MVPA at *T*2. However, despite the large mean differences in accelerometer-measured MVPA according to EVS classification, differences in minutes/week of accelerometer-measured MVPA did not reach statistical significance (*T*1 $P = 0.11$, *T*2 $P = 0.58$), likely due to inadequate power.

SBAS—Spearman's rank correlation coefficients between the SBAS and accelerometer-measured MVPA were 0.10 ($P = 0.59$) at *T*1 and 0.28 ($P = 0.15$) at *T*2. Sensitivity, specificity, and negative and positive predictive values were 18, 79, 33, and 62% at *T*1 and 67, 58, 43, and 79% at *T*2. Participants classified as meeting the 2008 PA Guidelines based on the SBAS engaged in 42.9 more min/week of accelerometer-measured MVPA than participants not meeting the guidelines at *T*1 and 19.3 more min/week of accelerometer-measured MVPA at *T*2. Similar to the findings of the EVS, differences in minutes/week of accelerometer-measured MVPA according to women classified as meeting the 2008 US PA Guidelines by the SBAS were not statistically different than those not meeting the guidelines (*T*1 $P = 0.27$, *T*2 $P = 0.28$).

Discussion

Limited research has evaluated the validity of brief PA questionnaires among AA women (Wolin *et al.*, 2010). The purpose of the current study was to assess validity of two brief PA questionnaires, the EVS and SBAS, against accelerometer-measured MVPA among a sample of highly educated AA women. Results provide important information regarding the practical application of each survey for use in both research and practice settings.

Validity

Both surveys demonstrated low-to-modest correlations with ActiGraph GT3X + measured MVPA (EVS $r = 0.27$ and $r = 0.26$ at T1 and T2, respectively; SBAS $r = 0.10$ and $r = 0.28$ T1 and T2, respectively). These correlations are comparable to results of other studies that have compared accelerometer-measured PA with questionnaires among various populations, including the IPAQ (Hagstromer *et al.*, 2006; Hallal *et al.*, 2010; Nicaise *et al.*, 2011), IPAQ-short version (Macfarlane *et al.*, 2007; Kurtze *et al.*, 2008), Global Physical Activity Questionnaire (Bull *et al.*, 2009; Trinh *et al.*, 2009), Epic Physical Activity Questionnaire (Cust *et al.*, 2008), and Modifiable Physical Activity Questionnaire (Jacobi *et al.*, 2009). Similar correlation coefficients have also been reported by among the few studies that have evaluated validity of PA questionnaires exclusively among AA. For example, Carter-Nolan *et al.* (2006) evaluated the PA questionnaire used in the Black Women's Health Study. Results showed a correlation of 0.28 between questionnaire and accelerometer-measured PA. Likewise, Smitherman *et al.* (2009) evaluated the validity of the Jackson Physical Activity Study among a sample of AA males and females. Results of their study showed a correlation of 0.24 between the PA questionnaire items and accelerometer-measured PA.

We reviewed the literature to identify studies that have examined the validity of the EVS and SBAS when compared with accelerometer-measured PA. Results of this review identified only one study examining validity of the SBAS against accelerometers and no studies evaluating the EVS against accelerometers. Vega-Lopez *et al.* (2014) conducted the sole study examining the validity of the SBAS against accelerometer-measured MVPA by examining the validity of a Spanish translated version of the SBAS among a sample of Mexican Americans. Results showed a correlation (Spearman rank) of 0.38 between the SBAS and accelerometer-measured PA, which is slightly higher than the correlations found in our study. Other validation studies of the SBAS have used subjective PA measures (ie, Stanford Seven-Day Recall) (Taylor-Piliae *et al.*, 2006), cardiovascular risk factors (Taylor-Piliae *et al.*, 2007), and/or psychosocial outcomes (Taylor-Piliae *et al.*, 2010) as criterion or discriminant measures of validity in samples from 500 to 1010 adults enrolled in the ADVANCE trial. Results of these studies have reported reasonable validity for the SBAS to differentiate PA levels, the prevalence of cardiovascular risk factors, and psychosocial outcomes among varying populations.

The initial validation of the EVS was published in 2012 and examined the face validity of the EVS against PA scores from two national population-based surveys, the 2005–06 National Health and Nutrition Examination Survey (NHANES) and the 2007 California Behavioral Risk Factor Surveillance System (BRFSS) (Coleman *et al.*, 2012). Results showed that PA levels assessed by the PA questionnaires associated with the NHANES and BRFSS surveys were higher than the PA levels arising from the EVS. These observations were similar to the findings in the current study for the pre-intervention assessment (T1) where PA levels were lower than those reported for PA measured by accelerometers (ie, accelerometer-measured PA = 128 min/week versus EVS measured PA = 95 min/week). Albeit the studies are few in number, the consistency of these findings suggest that the EVS may under-report PA in comparison with PA measured by both accelerometers and national population-based surveys. One explanation for the differences may be associated with the

fewer questions asked by the EVS ($n = 2$) in comparison with the national surveys (BRFSS n = 7; NHANES $n = 19$). Another possibility is that the small sample size in the current study ($n = 30$) may have contributed to the differences in the PA scores. It is important to note that we were unable to identify any studies examining the validity of the EVS against objective measures of PA. As such, the current study is a first step in exploring the criterion validity of the EVS with an objective PA measure.

We also observed a notable, although not statistically significant, difference in accelerometer-measured minutes/week of MVPA between participants classified as meeting the 2008 PA Guidelines versus those not meeting the guidelines for both PA questionnaires. At the pre-intervention assessment ($T1$), participants classified as meeting the 2008 PA Guidelines of 150 min of MVPA per week by the EVS accumulated 189 min/week of MVPA according to accelerometers versus 116 min/week for those classified as not meeting the guidelines ($P = 0.11$). Similarly, participants classified as meeting the national PA Guidelines by the SBAS performed 162 min/week of accelerometer-measured MPVA versus 119 min/week of MVPA among those classified as not meeting the PA Guidelines at the $T1$ assessment ($P = 0.27$). Comparable accelerometer-measured MVPA outcomes were found among participants classified as meeting versus not meeting the 2008 PA Guidelines for both questionnaires at $T2$; however, the difference in accelerometer-measured MVPA between those meeting and not meeting the guidelines according the SBAS was not as large at the $T2$ assessment (137 versus 118 min/week; $P = 0.28$) as compared with the EVS (194 versus 117 min/week; $P = 0.58$). While it is likely the small sample size was responsible for a failure to identify significant differences between activity classifications, the large standard deviations associated with the mean minutes per week of MVPA recorded by the accelerometer may be attributed to these non-significant outcomes.

Sensitivity, specificity, and positive and negative predictive values

Sensitivity and specificity of the SBAS were 18 and 79%, respectively, at the $T1$ assessment and 67 and 58% at the $T2$ assessment. These values are substantially lower than those reported in the original validation of the SBAS among a sample of older adults aged 60–69 years (Taylor-Piliae *et al.*, 2006), as well as in another validation study examining the validity of a Spanish translated version of the SBAS (Vega-Lopez *et al.*, 2014). The positive predictive values for the SBAS were also lower than those in previous validations studies, whereas, the negative predictive values were slightly higher (Taylor-Piliae *et al.*, 2006; Vega-Lopez *et al.*, 2014). The low sensitivity and low positive predictive values suggest that the SBAS has the potential to misclassify AA women who actually meet the 2008 PA Guidelines of 150 min/week of PA. On the other hand, the moderate-to-high specificity and negative predictive values of the SBAS indicate that the questionnaire is reasonably accurate for classifying AA women who do not meet the national PA Guidelines.In reference the EVS, sensitivity values were low (27% at $T1$, 33% at $T2$), while specificity values were fairly acceptable (89% at $T1$ and 74% at $T2$). Likewise, while we observed some variance in the positive predictive values for the EVS (59% at $T1$, 38% at $T2$); the negative predictive values were moderate and remained relatively stable at both assessment periods (68% at $T1$, 70% at $T2$). We reviewed the literature to identify other validation studies that have explored the sensitivity, specificity, and negative and positive predictive values for the EVS to

accurately classify individuals who meet and do not meet the National PA Guidelines; however, no such validation studies were identified. Thus, the findings of the current study provide important insight regarding the use of the EVS to accurately identify individuals meeting or not meeting the national PA Guidelines.

Research and practice implications of validation outcomes

The low sensitivity and positive predictive values for the EVS and SBAS indicate that both questionnaires are not well-suited to accurately identify AA women who achieve the national PA recommendations. Conversely, the moderate-to-high specificity and negative predictive values suggest that both questionnaires are reasonable measures to identify AA women who do not meet the national PA Guidelines. These outcomes provide practical implications for the use of both surveys in both research and practice settings. For example, the EVS and SBAS may be of interest for researchers who need a quick PA assessment method to screen potential participants for enrollment into PA studies, as many study protocols require participants to engage in less at 150 min of MVPA each week to meet eligibility criteria. Either questionnaire may also be advantageous for clinicians to use in health care settings in order to identify AA women who need to increase the PA levels.

The current validation evaluation also highlights instances where these surveys may not be ideal for use. Owing to the low sensitivity of both the EVS and SBAS in our sample of AA women, these questionnaires may not be useful to identify AA women who meet the national PA recommendations. For example, researchers interested in examining facilitators of PA among women who achieve the national PA Guidelines would not want to use the EVS or SBAS to assess PA in order to establish inclusion criteria for the study. Both surveys may have limited utility for researchers interested in evaluating changes in PA over time, as that is not what they were originally designed to do (Taylor-Piliae *et al.*, 2006; Coleman *et al.*, 2012). The EVS and SBAS also may not be ideal for use among AA women with low literacy and/or numeracy levels as the SBAS includes rather long questions and prompts in order to assess PA. Such formats may be difficult for AA women with low literacy levels to read and comprehend. Likewise, while the two questionnaire items of the EVS are much shorter than the prompts of the SBAS, numeracy may be an issue for some populations as the questionnaire requires participants to recall days of PA performed and average minutes per day PA was performed.

The current validation evaluation has several strengths. First, we used accelerometers as the criterion PA measure. Accelerometers are advantageous because they do not rely on participant recall; therefore increasing the validity of the PA assessment and limiting single source bias. Another strength of the study was that the study design allowed us to compare different PA measures simultaneously, which provided important insight on how each measure does when assessing PA over the same time period. Despite these strengths, several limitations must also be acknowledged. Eligibility criteria required participants to engage in <150 min/week of PA according to the IPAQ in order to enroll in the study. While participants were not explicitly informed of this criterion during recruitment and eligibility screenings, some may have under-report their PA to ensure they could participate in the study. Similarly, the sequence in which participants completed the PA questionnaires and

wore accelerometers could have influenced study outcomes. Participants completed the questionnaires before wearing an accelerometer. If data collection procedures were reversed and participants completed the questionnaires after wearing the accelerometer, outcomes may have differed (since the act of wearing an accelerometer can cause participants to alter normal PA patterns). However, we note that the underlying assumption of both self-report measures is that they assess an individual's 'usual' or 'regular' PA level (Taylor-Piliae *et al.*, 2006; Coleman *et al.*, 2012), which is similar to the assumption of using accelerometers to assess PA in free-living settings (Trost *et al.*, 2005). Therefore, we do not feel that the sequence of assessment had a major influence on the validity outcomes reported. The small sample size was also a limitation, as it limited statistical power and our ability to make any definitive conclusions regarding the validity of the PA measures. Finally, our sample was comprised of a group of fairly highly educated AA women residing in a metropolitan area of Phoenix, Arizona. Accordingly, generalization of our results to other subgroups of AA women should be done with caution.

Conclusions

Both the EVS and SBAS show promise for identifying AA women not meeting the 2008 PA Guidelines of 150 min of moderate-intensity PA each week. Results suggest that these surveys may be of particular use to identify AA women who need to increase their PA levels. Given the modest criterion validity against accelerometer-measured PA and low sensitivity values for the questionnaires, use of the surveys to identify AA women who meet the PA Guidelines should be done with caution. Few studies have evaluated the validity of self-report PA measures in AA women, and even fewer have evaluated brief PA questionnaires in this underserved population (Wolin *et al.*, 2010). Results of the current study provide important preliminary insight into the utility of both the EVS and SBAS to assess PA in AA women. Future studies with larger samples are warranted to further explore the validity of the EVS and SBAS among AA women.

Acknowledgment

References

Altman DG, Bland JM. Diagnostic tests 2: predictive values. British Medical Journal. 1994; 309:102. [PubMed: 8038641]

Ashraf MJ, Baweja P. Obesity: the 'huge' problem in cardiovascular diseases. Missouri Medicine. 2013; 110:499–504. [PubMed: 24564002]

Bastien M, Poirier P, Lemieux I, Despres JP. Overview of epidemiology and contribution of obesity to cardiovascular disease. Progress in Cardiovascular Diseases. 2014; 56(4):369–81. [PubMed: 24438728]

Bluher M. Adipose tissue dysfunction in obesity. Experimental and Clinical Endocrinology & Diabetes. 2009; 117:241–50. [PubMed: 19358089]

Bull FC, Maslin TS, Armstrong T. Global physical activity questionnaire (GPAQ): nine country reliability and validity study. Journal of Physical Activity and Health. 2009; 6:790–804. [PubMed: 20101923]

Carter-Nolan PL, Adams-Campbell LL, Makambi K, Lewis S, Palmer JR, Rosenberg L. Validation of physical activity instruments: black Women's Health Study. Ethnicity and Disease. 2006; 16:943–47. [PubMed: 17061751]

Centers for Disease Control and Prevention (CDC). Prevalence of fruit and vegetable consumption and physical activity by race/ethnicity-United States 2005. MMWR. Morbidity and Mortality Weekly Report. 2007; 56:301–4. [PubMed: 17410082]

Centers for Disease Control and Prevention (CDC). [15 March 2014] U.S. physical activity statistics. 2010. from http://www.cdc.gov/nccdphp/dnpa/physical/stats/index.htm

Coleman KJ, Ngor E, Reynolds K, Quinn VP, Koebnick C, Young DR, Sternfeld B, Sallis RE. Initial validation of an exercise "vital sign" in electronic medical records. Medicine and Science in Sports and Exercise. 2012; 44:2071–76. [PubMed: 22688832]

Cust AE, Smith BJ, Chau J, van der Ploeg HP, Friedenreich CM, Armstrong BK, Bauman A. Validity and repeatability of the EPIC physical activity questionnaire: a validation study using accelerometers as an objective measure. International Journal of Behavioral Nutrition and Physical Activity. 2008; 5:33. [PubMed: 18513450]

Ervin, RB. National Health Statistics Report. Vol. 13. U.S. Department of Health and Human Services, Centers for Disease Control and Preventions, National Center for Health Statistics; 2009. Prevalence of metabolic syndrome among adults 20 years of age and over, by sex, age, race and ethnicity, and body mass index: United States, 2003–2006.; p. 1-8.

Evenson KR, Rosamond WD, Cai J, Pereira MA, Ainsworth BE. Occupational physical activity in the atherosclerosis risk in communities study. Annals of Epidemiology. 2003; 13:351–57. [PubMed: 12821274]

Freedson PS, Melanson E, Sirard J. Calibration of the computer science and applications, Inc. accelerometer. Medicine and Science in Sports and Exercise. 1998; 30:777–81. [PubMed: 9588623]

Glauber H, Karnieli E. Preventing type 2 diabetes mellitus: a call for personalized intervention. Permanente Journal. 2013; 17:74–79. [PubMed: 24355893]

Goossens GH. The role of adipose tissue dysfunction in the pathogenesis of obesity-related insulin resistance. Physiology and Behavaior. 2008; 94:206–18.

Hagstromer M, Oja P, Sjostrom M. The International Physical Activity Questionnaire (IPAQ): a study of concurrent and construct validity. Public Health Nursing. 2006; 9:755–62.

Hallal PC, Simoes E, Reichert FF, Azevedo MR, Ramos LR, Pratt M, Brownson RC. Validity and reliability of the telephone-administered international physical activity questionnaire in Brazil. Journal of Physcial Activity and Health. 2010; 7:402–9.

He XZ, Baker DW. Differences in leisure-time, household, and work-related physical activity by race, ethnicity, and education. Journal of General Internal Medicine. 2005; 20:259–66. [PubMed: 15836530]

Jacobi D, Charles MA, Tafflet M, Lommez A, Borys JM, Oppert JM. Relationships of self-reported physical activity domains with accelerometry recordings in French adults. European Journal of Epidemiology. 2009; 24:171–79. [PubMed: 19283494]

Joseph RP, Ainsworth BE, Vega-Lopez S, Keller CS. Use of the Stanford brief activity survey for physical activity assessment in postpartum Latinas: a validation study of a linguistically translated Spanish version. Hispanic Health Care International. 2014; 12:146–54. [PubMed: 25239211]

Joseph RP, Keller C, Adams MA, Ainsworth BE. Print versus a culturally-relevant Facebook and text message delivered intervention to promote physical activity in African American women: a randomized pilot trial. BMC Women's Health. 2015; 15:1–18. [PubMed: 25608736]

Kahn SE, Hull RL, Utzschneider KM. Mechanisms linking obesity to insulin resistance and type 2 diabetes. Nature. 2006; 444:840–46. [PubMed: 17167471]

Kurtze N, Rangul V, Hustvedt BE. Reliability and validity of the International Physical Activity Questionnaire in the Nord-Trondelag health study (HUNT) population of men. BMC Medical Research Methodology. 2008; 8:63. [PubMed: 18844976]

Levine JA. Measurement of energy expenditure. Public Health Nutrition. 2005; 8:1123–132. [PubMed: 16277824]

Macfarlane DJ, Lee CC, Ho EY, Chan KL, Chan DT. Reliability and validity of the Chinese version of IPAQ (short, last 7 days). Journal of Science and Medicine in Sport. 2007; 10:45–51. [PubMed: 16807105]

Matthew CE. Calibration of accelerometer output for adults. Medicine and Science in Sports and Exercise. 2005; 37(11 Suppl):S512–522. [PubMed: 16294114]

Matthews CE, Chen KY, Freedson PS, Buchowski MS, Beech BM, Pate RR, Troiano RP. Amount of time spent in sedentary behaviors in the United States, 2003–2004. American Journal of Epidemiology. 2008; 167:875–81. [PubMed: 18303006]

Mosca L. Guidelines for prevention of cardiovascular disease in women: a summary of recommendations. Preventive Cardiology. 2007; 10(Suppl 4):19–25. [PubMed: 17917527]

Mosca L, Manson JE, Sutherland SE, Langer RD, Manolio T, Barrett-Connor E. Cardiovascular disease in women: a statement for healthcare professionals from the American Heart Association. Writing Group. Circulation. 1997; 96:2468–482. [PubMed: 9337227]

Nicaise V, Marshall S, Ainsworth BE. Domain-specific physical activity and self-report bias among low-income Latinas living in San Diego county. Journal of Physical Activity and Health. 2011; 8:881–90. [PubMed: 21885878]

Nies MA, Vollman M, Cook T. African American women's experiences with physical activity in their daily lives. Public Health Nursing. 1999; 16:23–31. [PubMed: 10074819]

Ogden CL, Carroll MD, Flegal KM. Prevalence of obesity in the United States. JAMA. 2014; 312:189–90. [PubMed: 25005661]

Pekmezi D, Marcus B, Meneses K, Baskin ML, Ard JD, Martin MY, Adams N, Robinson C, Demark-Wahnefried W. Developing an intervention to address physical activity barriers for African-American women in the deep south (USA). Women's Health. 2013; 9:301–12.

Plasqui G, Bonomi AG, Westerterp KR. Daily physical activity assessment with accelerometers: new insights and validation studies. Obesity Reviews. 2013; 14:451–62. [PubMed: 23398786]

Plasqui G, Westerterp KR. Physical activity assessment with accelerometers: an evaluation against doubly labeled water. Obesity. 2007; 15:2371–379. [PubMed: 17925461]

Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. Research Quarterly in Exercise and Sport. 2000; 17(2 Suppl):S1–14.

Smitherman TA, Dubbert PM, Grothe KB, Sung JH, Kendzor DE, Reis JP, Ainsworth BE, Newton RL Jr, Lesniak KT, Taylor HA Jr. Validation of the Jackson heart study physical activity survey in African Americans. Journal of Physcial Activity and Health. 2009; 6(Suppl 1):S124–132.

Taylor-Piliae RE, Fair JM, Haskell WL, Varady AN, Iribarren C, Hlatky MA, Go AS, Fortmann SP. Validation of the Stanford brief activity survey: examining psychological factors and physical activity levels in older adults. Journal of Physical Activity and Health. 2010; 7:87–94. [PubMed: 20231759]

Taylor-Piliae RE, Haskell WL, Iribarren C, Norton LC, Mahbouba MH, Fair JM, Hlatky MA, Go AS, Fortmann SP. Clinical utility of the Stanford brief activity survey in men and women with early-onset coronary artery disease. Journal of Cardiopulmonary Rehabilitation and Prevention. 2007; 27:227–32. [PubMed: 17667019]

Taylor-Piliae RE, Norton LC, Haskell WL, Mahbouda MH, Fair JM, Iribarren C, Hlatky MA, Go AS, Fortmann SP. Validation of a new brief physical activity survey among men and women aged 60–69 years. American Journal of Epidemiology. 2006; 164:598–606. [PubMed: 16840522]

Thomas S, Reading J, Shephard RJ. Revision of the Physical Activity Readiness Questionnaire (PAR-Q). Canadian Journal of Sports Sciences. 1992; 17:338–45.

Trinh OT, Nguyen ND, van der Ploeg HP, Dibley MJ, Bauman A. Test-retest repeatability and relative validity of the global physical activity questionnaire in a developing country context. Journal of Physical Activity and Health. 2009; 6(Suppl 1):S46–53. [PubMed: 19998849]

Trost SG, McIver KL, Pate RL. Conducting accelerometer-based activity assessments in field-based research. Medicine & Science in Sports & Exercise. 2005; 37(Suppl 11):S531–543. [PubMed: 16294116]

United States Census Bureau. Statistical abstract of the United States. 2014; 2012:151–152. from http://www.census.gov/prod/2011pubs/12statab/educ.pdf.

U.S. Department of Health and Human Services. 2008 physical activity guidelines for Americans. Department of Health and Human Services; USA: 2008.

Vega-Lopez S, Chavez A, Farr KJ, Ainsworth BE. Validity and reliability of two brief physical activity questionnaires among Spanish-speaking individuals of Mexican descent. BMC Research Notes. 2014; 7:29. [PubMed: 24410978]

Westerterp KR. Assessment of physical activity: a critical appraisal. European Journal of Applied Physiology. 2009; 105:823–28. [PubMed: 19205725]

Wolin KY, Fagin C, Ufere N, Tuchman H, Bennett GG. Physical activity in US Blacks: a systematic review and critical examination of self-report instruments. International Journal of Behavioral Nutrition and Physical Activitiy. 2010; 7:73.

Wood FG. Ethnic differences in exercise among adults with diabetes. Western Journal of Nursing Research. 2002; 24:502–15. [PubMed: 12148832]

Young DR, Jee SH, Appel LJ. A comparison of the Yale physical activity survey with other physical activity measures. Medicine in Science in Sports and Exercise. 2001; 33:955–61.
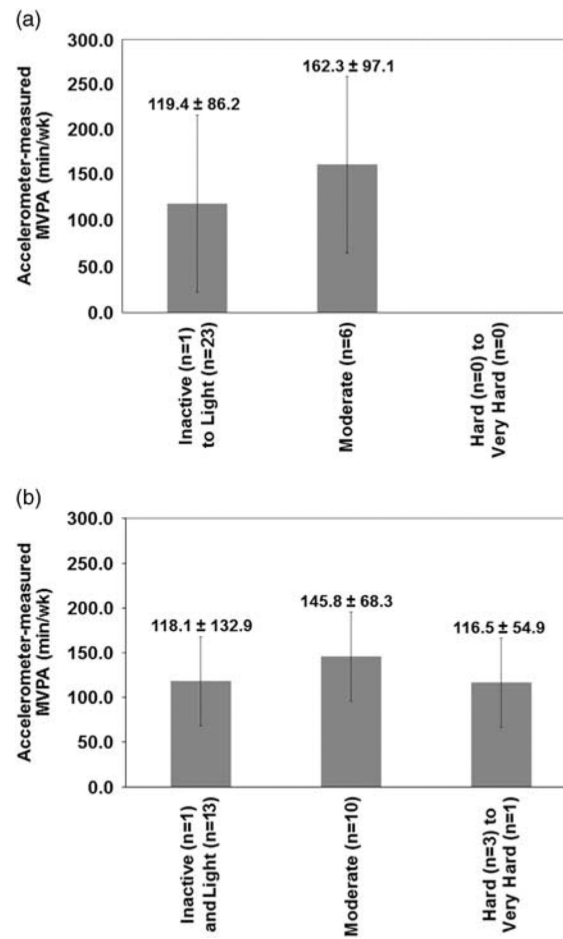
Figure 1.
Accelerometer-measured time spent in MVPA for participants classified in each PA category according to the SBAS. (a) SBAS category at pre-intervention assessment (*T*1). (b) SBAS category at post-intervention assessment (*T*2). MPVA = moderate-to-vigorous physical activity; PA = physical activity; SBAS = Stanford Brief Activity Survey

Table 1

Accelerometer-measured mean (SD) minutes/week of MVPA among participants classified as meeting or not meeting the 2008 PA Guidelines by PA assessment method

Time point	Accelerometer						Exercise vital sign						Stanford Brief Activity Survey					
	Meeting guidelines		Not meeting guidelines		p^a		Meeting guidelines		Not meeting guidelines		p^a		Meeting guidelines		Not meeting guidelines		p^a	
	n	Accelerometer measured MPVA (min/week)	n	Accelerometer measured MPVA (min/week)			n	Accelerometer measured MPVA (min/week)	n	Accelerometer measured MPVA (min/week)			n	Accelerometer measured MPVA (min/week)	n	Accelerometer measured MPVA (min/week)		
T1 (n = 30)	11	222.6 (66.6)	19	73.3 (38.9)	<0.001		5	189.0 (106.6)	25	115.8 (81.4)	0.11		6	162.3 (97.1)	24	119.4 (86.2)	0.27	
T2 (n = 28)	9	239.2 (107.0)	19	75.0 (40.4)	<0.001		4	193.8 (87.5)	24	116.8 (102.7)	0.58		14	137.4 (64.1)	14	118.1 (132.9)	0.28	

MPVA = moderate-to-vigorous physical activity; PA = physical activity; T1 = pre-intervention assessment; T2 = post-intervention assessment.

[a] P-value for independent t-test performed on the square root transformation of accelerometer-measured MVPA for participants meeting the 2008 US PA Guidelines compared with participants not meeting the guidelines.

Table 2

Validity, sensitivity, specificity, positive and negative predictive values of the Exercise Vital Sign and Stanford Brief Activity Survey to predict meeting the 2008 PA Guidelines according to accelerometer-measured MVPA

| | Exercise Vital Sign | | | | Stanford Brief Activity Survey | | | |
| | Pre-intervention (*T*1) | | Post-intervention (*T*2) | | Pre-intervention (*T*1) | | Post-intervention (*T*2) | |
	r	*P* value	*r*	*P* value	*r*	*P* value	*r*	*P* value
Validity[a]	0.27	0.15	0.26	0.17	0.10	0.59	0.28	0.15
Sensitivity[b]	27%		33%		18%		67%	
Specificity[c]	89%		74%		79%		58%	
Positive predictive value[d]	59%		38%		33%		43%	
Negative predictive value[e]	68%		70%		62%		79%	

MPVA = moderate-to-vigorous physical activity; PA = physical activity.

Positive and negative predictive values were adjusted for prevalence of participants meeting the 2008 PA Guidelines according to accelerometers based on formulas provided by Altman and Bland (1994).

[a] Spearman's rank ordered correlation between accelerometer-measured MVPA and score on each questionnaire.

[b] True positives/(true positives + false negatives).

[c] True negatives/(true negatives + false positives).

[d] (sensitivity × prevalence)/(sensitivity × prevalence + (1 – specificity) × (1 – prevalence)).

[e] (specificity × (1 – prevalence))/((1 – sensitivity) × revalence + specificity × (1 – prevalence)).