Title

Validation of the WHI Brief Physical Activity Questionnaire among Women Diagnosed with Breast Cancer

Journal

American Journal of Health Behavior, 31(2)

ISSN

1087-3244

Authors

Johnson-Kozlow, Marilyn
Rock, Cheryl L
Gilpin, Elizabeth A
et al.

Publication Date

2007

DOI

10.5993/ajhb.31.2.8

Peer reviewed

Validation of the WHI Brief Physical Activity Questionnaire among Women Diagnosed with Breast Cancer

Marilyn Johnson-Kozlow, PhD; Cheryl L. Rock, PhD, RD; Elizabeth A. Gilpin, MS
Kathryn A. Hollenbach, PhD; John P. Pierce, PhD

***Objective*: To investigate the psychometric properties of the physical activity (PA) measure of the Women's Health Initiative (WHI). *Methods*: Women diagnosed with breast cancer and enrolled in the Women's Healthy Eating and Living Study (average age 55 years) wore an accelerometer for 1 week and completed the 7-day Physical Activity Recall (PAR) and brief WHI measure. *Results*: Both self-reports correlated 0.73 with the accelerometer and had 100% sensitivity for meeting the current PA guideline, but specificity was significantly higher for the PAR. *Conclusions*: The WHI measure had comparable validity, sensitivity, and measurement bias compared to the widely accepted PAR.**

***Key words*: epidemiological measurements; exercise; psychometrics; questionnaires; sensitivity and specificity**

***Am J Health Behav.* 2007;31(2):193-202**

Considerable evidence suggests that physical inactivity may be causally associated with the development of several chronic diseases.[1,2] To reduce this risk, the Centers for Disease Control and Prevention (CDC) and the American College of Sports Medicine (ACSM) recommend that adults engage in 30 minutes or more of at least moderate physical activity on most days of the week.[3] Measurement of physical activity is thus a priority for large cohort studies and randomized trials that study the relationship between physical activity and chronic disease. To reduce participant burden and cost, many of these studies seek short, valid, self-administered measures of physical activity.

Several self-report physical activity surveys have been developed.[4] One commonly used instrument is the 7-Day Physical Activity Recall (PAR), a semistructured interview developed for the Stanford Five City Project.[5,6] However, this interview takes an average of 20 minutes to administer and requires interviewer training,[7] thus adding significant participant burden and cost to large-scale studies. Nevertheless, the PAR is a widely accepted measure of physical activity.[6] Several studies have demonstrated its validity and reported no significant difference in average energy expenditure using the doubly labeled water method.[8,9]

Marilyn Johnson-Kozlow, Research Scientist, San Diego State University Research Foundation, San Diego, CA. Cheryl L. Rock, Professor, Family and Preventive Medicine; Elizabeth A. Gilpin, Clinical Professor Emeritus, Family and Preventive Medicine; Kathryn A. Hollenbach, Associate Adjunct Professor, Family and Preventive Medicine; John P. Pierce, Professor, Family and Preventive Medicine, Cancer Prevention and Control Program, UCSD Cancer Center, La Jolla, CA.

Address correspondence to Dr Pierce, Cancer Prevention and Control Program, Moores UCSD Cancer Center, 3855 Health Sciences Drive, #0901, La Jolla, CA 92093-0901. E-mail: jppierce@ucsd.edu

The large-scale Women's Health Initiative (WHI) Clinical Trial and Observational Study[10] developed a brief, self-administered instrument to categorize the physical activity level of its participants. The WHI measure is a simple, 9-item, multiple-choice, paper-and-pencil questionnaire assessing recreational walking and light, moderate, and vigorous physical activities.[11] Test-retest reliability of this instrument, as evaluated by the weighted kappa statistic, at about 0.70, is moderate,[12] but its validity has not been evaluated yet.

Although there is consensus that the doubly labeled water technique is the best method for estimating total daily energy expenditure,[13] this method does not provide information about physical activity patterns and its high cost makes it impractical to use in most studies. A motion detector, such as an accelerometer,[14] is a relatively inexpensive alternative and commonly used method of objectively measuring physical activity.[6,15] This small instrument, which is worn on the body, measures vertical accelerations on a real-time basis and thus provides a measure of both duration and intensity of physical activity.[16,17] The basis for the use of the accelerometer as a measure of energy expenditure is that vertical acceleration is directly proportional to muscular forces and is therefore related to energy expenditure.[18,19]

In this investigation, we report data collected as an ancillary study to the Women's Healthy Eating and Living (WHEL) Study, a multisite randomized trial of a dietary intervention among women previously diagnosed with early stage breast cancer.[20] The WHEL Study obtained permission to use several of the self-report measures developed by the Women's Health Initiative, including their measure of physical activity. Under separate funding, this validation study was conducted among WHEL Study participants enrolled in the clinical site at the University of California, San Diego (UCSD). The measure developed by the Women's Health Initiative was validated by comparing it to an accelerometer as the criterion. Many physical activity self-report measures have been developed,[4,6] but few studies have compared the performance of different measures simultaneously.[21] When new measures are developed, it is particularly useful to compare their psychometric properties to well-established measures in the same sample. Although the WHI measure is shorter and self-administered and is thus less costly, it is not known whether it measures physical activity as accurately as the PAR. For this reason, because the 7-Day Physical Activity Recall is an accepted standard of physical activity measurement, it is also included in order to compare the validity obtained on the as yet unvalidated measure developed by the Women's Health Initiative.

METHODS

Procedures

A convenience sample of 74 women was recruited from the 531 women enrolled in the WHEL Study clinical site at the University of California, San Diego. Between October 2001 and July 2002, while attending their regularly scheduled clinic visit, WHEL Study participants were invited by WHEL Study staff to enroll in the physical activity measurement study. Height and weight were measured at the clinic visit by the UCSD General Clinical Research Center nursing staff, from which body mass index (weight (kg) / height $(m)^2$) was determined. Following the clinic visit, participants who expressed interest completed a half-hour orientation session with the first author, during which written informed consent was obtained. Study orientation included an introduction to and a demonstration of the appropriate way to wear the accelerometer; all women were individually fitted with the accelerometer. Women were given written instructions for wearing the accelerometer, an envelope to return the accelerometer, the WHI physical activity questionnaire to be completed at the end of 7 days, and a calendar showing when to wear the accelerometer. A copy of instructions and all study materials are available for review.[22] The PAR telephone interview was also scheduled to occur at the end of the 7-day period.

During the telephone interview but prior to administration of the PAR, participants were asked whether they had performed any of 25 physical activities listed on the National Health Interview Survey.[23] The purpose of this short survey was to obtain data about the types of activities the women did during the monitoring week. Women were also asked at this time whether there was anything

about wearing the accelerometer they did not like.

Measures

The Actigraph accelerometer. The accelerometer, Actigraph Model 7164 (formerly the Computer Science and Application [CSA], now manufactured by Manufacturing Technology Inc, Fort Walton Beach, Fla) was worn on a belt fitted around each woman's waist in a pouch midway between the navel and the right hip. Participants were asked to wear the accelerometer at all times, except when they retired to bed or did water-based activities such as swimming, because the accelerometer is not waterproof. The accelerometer was programmed to record vertical accelerations in the standard 1-minute intervals. After the accelerometers were returned, the data were downloaded using the manufacturer's computer software, which provided time spent in moderate and vigorous physical activity. Moderate physical activity corresponded to 3.0 to 5.9 METs (or 1953 to 5724 activity counts), where 1 MET, or metabolic equivalent, is defined as energy expended in comparison to resting metabolism during quiet sitting; and vigorous physical activity corresponded to 6.0 METs or more (>5724 activity counts), based on published calibration data.[16] Nonadherence to wearing the accelerometer was defined as no recorded movement for 8 or more hours on a scheduled monitoring day. Based on accelerometer data, the physical activity recommendation[3] was met if 30 or more minutes were accumulated per day for 5 or more days during the monitoring week. Thirty minutes could be accumulated in 3 10-minute bouts, two 15-minute bouts, or one bout of 30 minutes of physical activity. A report of each woman's physical activity based on the accelerometer was distributed as an incentive for participation.

WHI measure. The WHI measure questionnaire was composed of 9 multiple-choice items assessing recreational walking and light, moderate, and vigorous physical activity using a standard frequency by duration item format. Written directions on the questionnaire[22] instructed women to report only physical activities they performed while wearing the accelerometer. The directions of the WHI questionnaire were modified slightly in the present study to reflect a shorter time frame: physical activity was assessed during the past 7 days, instead of usual physical activity over the past month. There were 3 items assessing walking (pace, duration ["How many minutes did you usually walk?"], and frequency ["How often did you walk outside...?"]) and 2 items each assessing duration ("How long did you usually do...exercise like this at one time?") and frequency ("How often ... did you do the following exercises?]) of moderate and vigorous physical activity. There were an additional 2 items assessing duration and frequency of light physical activity, but following CDC/ACSM recommendations targeting the health effects of at least moderate physical activity,[3] these items were not analyzed. Pace of walking ("What was your usual walking speed?") had 4 categories: less than 2 mi/hr, 2 - 3 mi/hr, 3 - 4 mi/hr, and more than 4 mi/hr. The questionnaire used a 5-point scale (never, 1 time, 2 - 3 times, 4 - 6 times and 7+ times) to assess the frequency of walking outside the home and the frequency of other moderate and vigorous physical activity (1 day, 2 days, 3 days, 4 days, or 5+ days). Duration of walking, moderate, and vigorous physical activity was assessed on a 4-point scale (less than 20 minutes, 20 - 39 minutes, 40 - 59 minutes, and 1 hour or more). Exemplars (or examples of types of physical activities) included in the items measuring moderate physical activity were those that were "not exhausting," such as biking outdoors, workouts with an exercise machine, calisthenics, easy swimming, and popular or folk dancing. Exemplars of vigorous physical activity, described as those that made the participant "work up a sweat and [her] heart beat fast" were aerobics, aerobic dancing, jogging, tennis, or lap swimming.

To obtain time spent in physical activity per week, frequency and duration multiple-choice categories were multiplied, which were recoded to their midpoint value. Frequency of walking was recoded as 0 times per week for rarely or never, 1 for 1, 2.5 for 2 to 3, 5 for 4 to 6, and 7 for 7 or more times per week. Five or more days of moderate and vigorous physical activity was recoded as 5 times per week. Duration of walking, moderate and vigorous physical activity less than 20 min was recoded as 15, 20 to 39 as 30, 40 to 59 as 50, and 1 hour or more as 60 min/

week. Moderate physical activity was the sum of moderate exercise and walking at a pace of 2 - 4 mph. Vigorous physical activity was the sum of vigorous exercise and walking at a pace of more than 4 mph. Based on the WHI measure, the physical activity recommendation[3] was met if the number of days doing moderate and vigorous physical activity was 5 or more and the duration was 30 minutes or more per bout.

Seven-day Physical Activity Recall (PAR). The PAR was administered by the first author, previously trained to criterion in the standard protocol according to published evaluative criteria.[7] Briefly, participants were asked to estimate the time during each of the previous 7 days that was spent in moderate, hard, and very hard intensity physical activity. Time spent in each activity as a function of its intensity is recorded in bouts of quarter-hours: .25, .50, .75, and so forth, according to a rounding algorithm (ie, 10 – 22 min activity was recorded as .25 hour). This approach uses cues and prompts to help participants recall significant events over the previous 7 days, starting with the most recent day and working backward in time. The interviewer then uses these events to frame participants' recall of physical activity. Participants classified each physical activity as moderate, hard, and very hard by using walking at a brisk pace as the standard for moderate activity and running for very hard physical activity. Participants reported only physical activity they did while they wore the accelerometer. Total physical activity was based on the sum of time spent in moderate and vigorous activities. Based on the PAR, the physical activity recommendation[3] was met if the number of days doing at least moderate physical activity was 5 or more and the duration was 30 minute or more per day.

Statistical Analysis

All analyses were conducted using standard statistical software (SPSS, Inc., Chicago Ill, 1999). A 2-tailed hypothesis was used for all statistical analyses with an alpha level set at 0.05. Manually input data were evaluated for the presence of inaccurately input data and cleaned accordingly; downloaded data from the accelerometers were evaluated for data accuracy as well. Based on the Kolmogorov-Smirnov statistic, none of the measures of physical activity were normally distributed. Measurement bias was defined as the difference in self-reported minutes minus accelerometer minutes per week of physical activity. The Wilcoxon signed-rank test was used to compare the differences between continuous scores on self-report and accelerometer measures. Criterion-related validity of the 2 self-report physical activity instruments was assessed by Spearman rank-order correlation with the accelerometer score. The difference between the correlation coefficients was compared using a paired-sample t-test developed by Williams, which takes into account the correlation between the self-report scores.[24] For all measures, following CDC/ACSM recommendations,[3] total physical activity was limited to the sum of time spent in moderate and vigorous physical activity.

The proportion of women meeting the CDC/ACSM physical activity guideline (of 30 minutes or more of at least moderate activity on 5 or more days a week) from each measure was assessed. Because the guideline is on a per week basis, only those who wore the accelerometer all 7 days of the monitoring week (n = 58) were included in this analysis. McNemar's test for correlated proportions determined whether the proportions differed significantly from those obtained from the accelerometer. Sensitivity and specificity statistics and the 95% confidence interval around their difference were calculated for the 2 self-report measures. Bland-Altman plots,[25] intraclass correlations, and weighted kappa statistics were analyzed to indicate the agreement of WHI and PAR scores and classifications with accelerometer-based data.

RESULTS

The Study Population

Of the 74 women enrolled in the study, 11 were excluded from the analysis. One was excluded because of an accelerometer malfunction; 8 had more than half of their physical activity, such as water-based activity, not measured by the accelerometer; and 2 others did not complete the PAR interview. Although when asked, over half of the women reported problems or discomfort wearing the accelerometer, the remaining 63 participants wore the accelerometer on average 15.4 (SD = 1.5) hours per day; this was estimated to be 98% of their waking hours. Most women

Table 1
Accelerometer Descriptive Statistics and Difference With Self-Report Measures in Minutes/Week (n = 63)

	Descriptive Statistics: Accelerometer			Self-report Minus Accelerometer: WHI Measure		Self-report Minus Accelerometer: 7-Day PAR	
	Mean (SD)	Median	Range	$Mean_{dif}$ (SD_{dif})	$Range_{dif}$	$Mean_{dif}$ (SD_{dif})	$Range_{dif}$
Moderate	154 (94)	143	(12, 395)	-21 (94)*	(-203, 246)	+17 (102)	(-203, 240)
Vigorous	10 (35)	0	(0, 240)	+27 (60)**	(-82, 250)	+5 (32)	(-97, 162)
Total	165 (106)	159	(12, 443)	+6 (103)	(-193, 245)	+22 (104)	(-188, 240)

Note.
Total is moderate plus vigorous physical activity.
*** P<0.05. ** P<0.01; P values indicate a difference between self-report and accelerometer scores based on a 2-sided Wilcoxon signed-rank test.**

(60%) wore the monitor every hour during the week;16% did not wear the monitor for 1 hour at most during the week, and 14% did not wear the monitor for 2 to 7 hours during the week. Of the 63 participants, 58 wore the accelerometer for 7 full days, and this was the sample on which screening statistics were computed.

Study participants ranged in age from 38 to 72 years, with an average age of 55 years (SD = 9.1). Body mass index varied widely from 19 through 46 with a mean of 26.9 kg/m^2 (SD = 5.8). The majority of participants were college-educated (56%) non-Hispanic whites (87%). The most commonly reported physical activities were walking for exercise (84%), housework (82%), calisthenics/general exercise (52%), gardening (35%), and weight lifting (33%). Using the accelerometer data, women spent on average 97.4% of their waking hours in light physical activity, 2.4% in moderate physical activity, and 0.2% in vigorous physical activity.

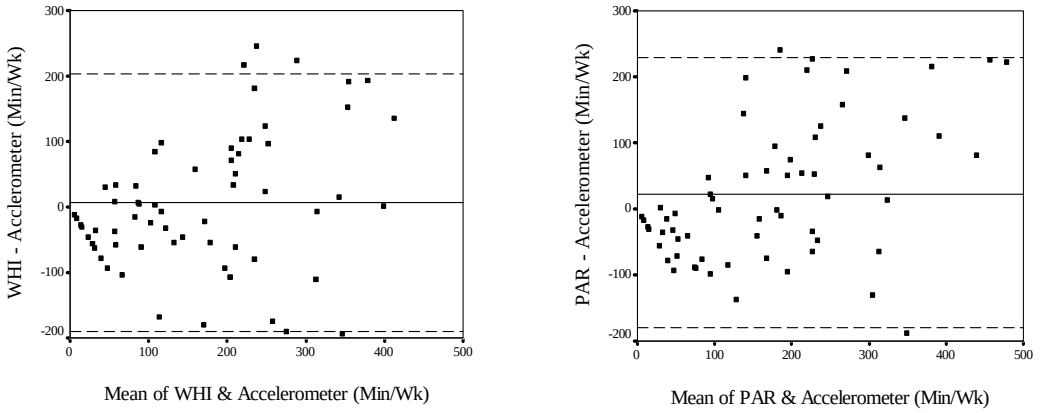
Comparison of Measures

Mean total (moderate plus vigorous) physical activity in this study population was 165 minutes per week (SD = 106) from the accelerometer, 187 (SD = 153) from the PAR, and 171 (SD = 137) from the WHI measure, corresponding to 24, 27, and 24 min/day, respectively (Table 1). On the PAR, individuals averaged 17 minutes of moderate and 5 minutes of vigorous physical activity more than indicated by the accelerometer (Table 1). These 22 (or 13%) additional minutes measured on the PAR were not statistically significantly different from the accelerometer estimate. However, there was a wide range in differences between the 2 measures (varying from –188 minutes to +240 minutes per week in total physical activity).

Compared to the accelerometer, the WHI measure did not provide a significantly different estimate of total physical activity (+6 minutes or 4%, P=0.95; Table 1). However, this small disparity masked significant variation in time spent in moderate vs vigorous physical activity. The WHI measure underestimated moderate activity, compared to the accelerometer (-21 minutes, P=0.08), but overestimated the amount of vigorous activity (+27 minutes, P<0.001). Similar to the PAR, there was a wide range in differences between the 2 measures, varying from -193 to +245 minutes per week in total physical activity.

Figure 1 shows the Bland-Altman plots for the 2 self-report measures of total physical activity. In order to assist in interpretation, a standardized scale was used for both measures. The solid line on the plot indicates the mean of the difference between self-report and accelerometer score. These values are the same as those reported previously: +6 and +22 min/wk for the WHI and PAR measures, respectively (Table 1). The dashed lines indicate the 95% confidence interval (CI) around the mean of the difference scores between self-report and accelerometer-based total physical activity. Outliers were those outside the 95% CI: there were comparable numbers of outliers in both the PAR and WHI measure. Based on

Figure 1
Bland-Altman Plots of Self-reported Total Physical Activity (n = 63)



a visual examination of their scatterplots, both instruments displayed about the same level of agreement between self-report and accelerometer-based scores. The fan-shaped scatter indicated that as physical activity increased, so did the overestimation on the self-report measures; this was true of both the WHI and PAR measures.

Both self-report measures of total physical activity were significantly and identically correlated with the accelerometer measure (r = 0.73, P<0.01). The correlation of total physical activity between the 2 self-report measures was 0.88 (P<0.001). The intraclass correlation for the PAR and WHI measures were also comparable at 0.68 (95% CI = 0.53-0.80) and 0.64 (95% CI = 0.47 – 0.78), respectively.

Meeting the CDC/ACSM Physical Activity Guideline

Only 6 women (10% or 6 of 58 women) met the current physical activity recommendation using the accelerometer estimate of physical activity (Table 2). The PAR interview estimated that 16 women (28%, or 16 of 58 women) met this criterion, whereas the WHI measure estimated that 27 women (47%, or 27 of 58 women) met this criterion. Thus, both of the self-report measures resulted in statistically significant overestimates of the proportion of women who met the recommended guideline (P<0.01) compared to the accelerometer.

All 6 women who were categorized as meeting the recommended guideline based on the accelerometer were also classified as meeting the guideline by both self-report measures, resulting in a sensitivity of 100%. The specificity for each of the measures, or the proportion correctly classified as not meeting the guideline, was 0.81 (or 42 of 52 women) for the PAR interview and 0.60 (or 31 of 52 women) for the WHI measure (Table 2). The PAR specificity was significantly greater (P=0.003); the 95% confidence interval around the difference in the specificities varied from 0.13 and 0.29. Based on Cohen's weighted kappa, there was fair-to-good agreement for the PAR (K = 0.46) and poor agreement for the WHI measure (K = 0.23), using published interpretation guidelines[26] (Table 2).

DISCUSSION

Both the more intensive 7-Day Physical Activity Recall and the brief WHI measure had equivalent Spearman correlations and comparable intraclass correlations with estimated total physical activity from the objective accelerometer measure. These results suggest that the WHI measure may meet both the reliability[12] and validity requirements for use in studies involving participants similar to those of the WHEL Study. All 3 measures of physical activity indicated that the highly

Table 2
Number of Participants Meeting Physical Activity Guideline by Self-Report Measure vs Accelerometer (n = 58)[a]

	Accelerometer			Screening Statistics		Agreement	
Self-report	Met guideline	Did not meet	P[b]	Sensitivity	Specificity[c]	(%)	Kappa
WHI Measure							
Met guideline	6	21	< 0.001	1.00	0.60	64	.23
Did not meet	0	31					
7-Day PAR							
Met guideline	6	10	0.002	1.00	0.81	83	.46
Did not meet	0	42					

Note.
a The 58 women who wore the accelerometer for 7 full days were used for this analysis, which determined whether they met the physical activity guideline of doing 30 minutes or more of at least moderate physical activity on 5 or more days during a 7-day week.
b The P values are based on a 2-sided test of McNemar's Test for Correlated Proportions, which tests whether the proportions on the self-report measure differ from the accelerometer.
c Based on McNemar's test for correlated proportions, the specificities for the 2 self-report measures were significantly different from each other at P = 0.003.

educated, middle-aged women in this study accumulated on average about 25 minutes a day of at least moderate activity, suggesting that the absolute validity[6] of the measures was comparable. The results indicated that although there was considerable within-individual variability among the 3 measures, the estimates of total physical activity obtained from each instrument were highly correlated (r = 0.88; intraclass correlation = 0.83).

The correspondence between measures was not maintained when total physical activity was separated into moderate vs vigorous activity. Although PAR estimates of both moderate and vigorous activity did not differ significantly from the accelerometer, the WHI measure significantly overestimated vigorous and underestimated moderate physical activity. The underestimation of moderate physical activity is not surprising because unlike the PAR, the WHI instrument does not assess gardening and housework, 2 of the more common physical activities performed by women in this study. Thus, the WHI instrument may be more likely to underestimate the health benefits of moderate and overestimate the benefits of vigorous physical activity.

Similarly, the choice of physical activity instrument affects the estimate of the absolute amount of physical activity individuals engage in compared to the CDC/ACSM-recommended level.[3] Although both self-report measures had 100% sensitivity, the PAR had a significantly higher specificity at 0.83 compared to 0.60 for the WHI measure. As all of the error was from classifying participants as meeting the guideline when the objective measure indicated otherwise, the self-report instruments, particularly the WHI instrument, would inaccurately ascribe health benefits to those so misclassified. The comparative difference in specificities between the instruments was reflected in their kappa coefficients, which were 0.46 vs. 0.23, for the PAR and WHI measures, respectively.

It may be expected that the PAR would provide the more accurate estimate of physical activity, because it maximizes recall in a number of well-researched ways. For example, the PAR uses a guided memory technique in which interviewees are encouraged to visualize the location of a past event to recall event-specific details. This technique has been shown to improve recalls of dietary behaviors[27,28] and physical activity.[29-31] Studies suggest that providing cues in this way may help

elicit recall of commonly performed incidental physical activities, such as walking.[29,31,32] Further, the PAR measures physical activity in all domains of life (occupational, recreational, transportation related, and house and yard work); whereas the WHI measure focused primarily on recreational activities. Although the estimate from the PAR was a nonsignificant 22 minutes higher than the estimate obtained from the accelerometer, the accelerometer does not measure water-based activity or physical activity done on an incline or above the waist.[17] The validity of the PAR found in the present investigation was somewhat higher than previous[4,6] validity estimates when using an accelerometer as the criterion: observed validity correlation coefficients were 0.50, 0.53, and 0.57.

Because this study used a small convenience sample and focused on a single week in the lives of well-educated women who were participants in a diet intervention trial, the findings may not generalize beyond this population subgroup. Additionally, the optimal design for this study would involve repeated measures on multiple instruments. Although this would have significantly strengthened the study's findings, this design was not possible within our resource and time constraints. We altered the questions referring to walking on the WHI measure from a per month basis to a per week basis in order to fit the design of the present investigation and used recoding to estimate continuous values on the WHI measure. Because item wording and response formatting may affect psychometric properties of self-report items,[33] these item modifications may limit conclusions regarding validity and measurement bias of this instrument. The WHI measure is self-administered whereas the PAR is interview administered, making the instruments not strictly comparable. An accelerometer was used to validate the self-report measures because it provided an objective measure of physical activity that was feasible and cost-effective. However, the monitor has a number of limitations. As stated previously, the monitor underestimates several sources of physical activity[17,19] and does not measure water-based activities. On the other hand, the technical reliability of the Actigraph indicates that it is very precise.[34] Adherence to wearing the accelerometer was good; however, 24% of the women did not wear the monitor for 2 or more hours during the monitoring week. When the nonadherent women were deleted from the sample, however, the validity correlation coefficients of the self-report measures with and without the nonadherent women were comparable (N = 48; r = 0.72 and 0.74, for the PAR and WHI measure, respectively).

The amount of variability in these measures suggests that a more stable estimate would be obtained by combining 2 or more measures of physical activity. This may be accomplished by administering the same measure multiple times, combining independent measures assessing different aspects of physical activity, or using different methods of physical activity assessment. In the future, researchers studying the validity of the WHI measure might consider including housework and gardening exemplars in the item set assessing moderate physical activity to determine whether this would counteract underestimation of moderate physical activity.

Although study participants were diagnosed with breast cancer, most (79%) indicated that their physical activity was not limited in any way during the week that they wore the monitor. Of the 21% who reported that their physical activity was limited, most (50%) of the limitations were due to a cold, migraine, or sinus infection, 35% were due to musculoskeletal problems (ie, knees and feet), and 15% were disease-related (ie, anemia and lymphedema). These results suggest that most of the women in our sample, who on average were 4.7 years postdiagnosis, appeared to be in fair to good health. Our data indicated that 87% of our sample walked for fun, transportation, or exercise. By comparison, 77% of women aged 38-72 in San Diego County walked for fun, transportation, or exercise, according to the 2003 California Health Interview Survey (http://www.chis.ucla.edu/). These results suggest that the women who participated in the present study were at least as physically active as San Diegan women, at least with respect to walking, the most common of all physical activities.

In summary, this validation study found support for using the WHI measure among women similar to those enrolled in the WHEL Study. However, the results should

be treated with caution as they are based on a small sample of well-educated women. In contrast to the PAR, which must be administered by a trained interviewer, the WHI measure is self-administered. Although the PAR requires about 20 minutes to administer, participants can complete the paper-and-pencil instrument in 5 minutes or less, indicating superior feasibility and cost-effectiveness. If total physical activity is the outcome of interest, the WHI measure performed well in terms of validity, measurement bias, and sensitivity. However, these results should be replicated in a larger, more diversified sample in order to improve the generalizability of the findings.

Acknowledgments

This study was supported by a National Cancer Institute grant #CA 72092 entitled "Behavioral Research in Cancer Prevention." It was approved as an ancillary study by the principal investigators of the Women's Healthy Eating and Living (WHEL) Study grant CA 69375 from the National Cancer Institute. The authors gratefully acknowledge the cooperation of the University of California, San Diego (UCSD) General Clinical Research Center nurses and the WHEL Study clinical site at UCSD, especially Dr Vicky Jones, Co-principal Investigator, and Dr Michelle McKinney, Site Coordinator. ■

REFERENCES

1. Powell KE, Thompson PD, Caspersen CJ, et al. Physical activity and the incidence of coronary heart disease. *Ann Rev Pub Health.* 1987;8:253-287.
2. U.S. Department of Health and Human Services. Healthy People 2010. 2nd ed. Understanding and Improving Health and Objectives for Improving Health. 2 Vols. Washington, DC: U.S. Government Printing Office, November 2000.
3. Pate RR, Pratt M, Blair SN, et al. Physical activity and public health: a recommendation from the Centers for Disease Control and Prevention and the American College of Sports Medicine. *J Am Med Assoc.* 1995;273:402-407.
4. Kriska AM, Caspersen CJ, eds. A collection of physical activity questionnaires for health-related research. *Med Sci Sports Exerc.* 1997;29(Suppl. 6):S89-S103.
5. Blair SN. How to assess exercise habits and physical fitness. In: Matarazzo JD, Herd JA, Miller NE, et al., eds. Behavioral Health: A Handbook of Health Enhancement and Disease Prevention. New York, NY: Wiley, 1984:424-447.
6. Sallis JF, Saelens BE. Assessment of physical activity by self-report: status, limitations, and future directions. *Res Q Exerc Sport.* 2000;71:1-14.
7. Sarkin JA, Campbell L, Gross L, et al. Project GRAD Seven-Day Physical Activity Recall Interviewer's Manual. *Med Sci Sports Exerc.* 1997;29 (Suppl 6): S91-S102.
8. Leenders NYJM, Sherman WM, Nagaraja HN, et al. Evaluation of methods to assess physical activity in free-living conditions. *Med Sci Sports Exerc.* 2001;33:1233-1240.
9. Racette SB, Schoeller DA, Kushner RF. Comparison of heart rate and physical activity recall with doubly labeled water in obese women. *Med Sci Sports Exerc.* 1995;27:126-133.
10. Women's Health Initiative Study Group. Design of the Women's Health Initiative Clinical Trial and Observational Study. *Control Clin Trials.* 1998;19:61-109.
11. McTiernan A, Kooperberg C, White E, et al. Recreational physical activity and the risk of breast cancer in postmenopausal women: the Women's Health Initiative Cohort Study. *J Am Med Assoc.* 2003;290:1331-1336.
12. Manson JE, Greenland P, LaCroix AZ, et al. Walking compared with vigorous exercise for the prevention of cardiovascular events in women. *N Engl J Med.* 2002;347:716-25.
13. Montoye HJ, Kemper HCG, Saris WHM, et al. Measuring Physical Activity and Energy Expenditure. Champaign, IL: Human Kinetics, 1996.
14. Bouten CVC, Sauren AAHJ, Verduin M, et al. Effects of body placement and orientation of body-fixed accelerometers on the assessment of energy expenditure during walking. *Med Biol Eng Computing.* 1997;35:50-56.
15. Freedson PS, Miller K. Objective monitoring of physical activity using motion sensors and heart rate. *Res Q Exerc Sport.* 2000;71:21-29.
16. Freedson PS, Melanson E, Sirard J. Calibration of the Computer Science and Applications, Inc. accelerometer. *Med Sci Sports Exerc.* 1998;30:777-7781.
17. Welk GJ, Blair SN, Wood K, et al. A comparative evaluation of 3 accelerometry-based physical activity monitors. *Med Sci Sports Exerc.* 2000;32 (Suppl 9):S489-497.
18. Melanson EL, Freedson PS. Validity of the Computer Science and Applications, Inc. (CSA) activity monitor. *Med Sci Sports Exerc.* 1995:27:934-940.
19. Hendelman D, Miller K, Baggett C, et al. Validity of accelerometry for the assessment of moderate intensity physical activity in the field. *Med Sci Sports Exerc.* 2000;32 (Suppl 9):S442-449.
20. Pierce JP, Faerber S, Wright FA, et al. for the Women's Healthy Eating and Living (WHEL) Study Group. A randomized trial of the effect of a plant-based dietary pattern on additional breast cancer events and survival: the Women's Healthy Eating and Living (WHEL) Study.

Control Clin Trials. 2002;23:728-756.
21. Jacobs DR Jr, Ainsworth BE, Hartman TJ, Leon AS. A simultaneous evaluation of 10 commonly used physical activity questionnaires. *Med Sci Sports Exerc.* 1993;25:81-91.
22. Johnson-Kozlow MF. Validity and measurement bias in 3 self-report measures of physical activity among women diagnosed with breast cancer [dissertation]. San Diego, CA: University of California, San Diego, 2003. ISBN 0-496-47520-7. DAI-B 64/07, p 3223, January 2004.
23. U.S. Department of Health and Human Services. Questionnaires from the National Health Interview Survey, 1985-1989. Vital and Health Statistics, Series 1, No.31, 1993.
24. Steiger JH. Tests for comparing elements of a correlation matrix. *Psychol Bull.* 1980;87:245-251.
25. Bland JM, Altman DJ. Statistical methods for assessing agreement between 2 methods of clinical measurements. *Lancet.* 1986;1:307-310.
26. Fleiss JL. Statistical Methods for Rates and Proportions. 2nd ed. New York: John Wiley and Sons; 1981.
27. Fisher RP, Quigly KL. Applying cognitive theory in public health investigations: enhancing food recall with the cognitive interview. In: Tanur JM, ed. Questions About Questions: Inquiries into the Cognitive Bases of Surveys. New York, NY: Russell Sage Foundation, 1992:154-169.
28. Tucker C, Bennett C. Procedural effects in the collection of consumer expenditure information: the diary operation test. Proceedings of the Section on Survey Research Methods. Washington, DC: American Statistical Association, 1988:256-261.
29. Durante R, Ainsworth BE. The recall of physical activity: using a cognitive model of the question-answering process. *Med Sci Sports Exerc.* 1996;28:1282-1291.
30. Swan PD, Spitler DL, Whiddon S, et al. Assessment of lifetime patterns of dairy food intake and physical activity. *Nutr Health.* 1997;4:255-270.
31. Johnson MF, Sallis JF, Hovell MF. Self-report assessment of walking: effects of aided recall instructions and item order. *Meas Phys Educ Exerc Sci.* 2000;4:141-155.
32. Menon G. Are the parts better than the whole? The effects of decompositional questions on judgments of frequent behaviors. *J Market Res.* 1997;34:335-346.
33. Schwarz N. Self-reports: how the questions shape the answers. *Am Psychol.* 1999;54:93-105.
34. Metcalf BS, Curnow JSH, Evans C, et al. Technical reliability of the CSA activity monitor: the EarlyBird Study. *Med Sci Sports Exerc.* 2002;34:1533-1537.