The purpose of this study was to determine the construct-related validity of self-reported historical walking, running, and jogging (WRJ) activity on the basis of data from the Aerobics Center Longitudinal Study (Dallas, Texas). A total of 4,100 men and 963 women underwent at least one medical examination between 1976 and 1985 and completed a follow-up questionnaire in 1986. Levels of glucose, cholesterol, and triglycerides, resting systolic blood pressure, body mass index (weight (kg)/height (m)^2), and cardiorespiratory fitness were measured at the time of the medical examination. The follow-up questionnaire assessed WRJ and other strenuous activities for each year from 1976 through 1985. Data analysis included Spearman and partial correlations, analysis of variance, analysis of covariance, and t tests. Results indicated significant correlations between recalled WRJ and treadmill times for each year throughout the 10-year period (r = 0.40–0.61). Participants were classified as historically either sufficiently physically active to receive a health benefit or insufficiently active for a health benefit. Engaging in sufficient levels of historical WRJ was associated with higher treadmill times and lower body mass indices for men and women and lower triglyceride levels for men. Self-reported historical WRJ can be assessed with reasonable validity in comparison with measured treadmill performance, with no decay in accuracy of reporting for up to 10 years in the past.

exercise; mental recall; motor activity; physical fitness; reproducibility of results

Abbreviations: MET, metabolic equivalent; WRJ, walking, running, and jogging.

The benefits of participation in regular physical activity are well known and widely documented (1). Engaging in physical activity has been associated with a reduced risk of specific chronic diseases such as cancer (2–6), cardiovascular disease (7, 8), and diabetes mellitus (9), as well as reduced risk of all-cause mortality (10–12). Physical activity is also directly related to obesity and weight control through energy balance (13) and has been found to be protective against some obesity-related problems, including gallstones, osteoarthritis, sleep apnea, and lower health-related quality of life due to days of poor health (14).

Most epidemiologic studies on the association between physical activity and health have measured current levels of physical activity; few studies have assessed self-reported information on historical physical activity. Historical physical activity is activity engaged in more than 1 year prior to the time of assessment (15–17). Valid methods of assessing historical physical activity may be useful in chronic disease epidemiology because, for some research questions, lifetime or long-term patterns of physical activity may help clarify associations between activity and health outcomes. In addition, historical physical activity questionnaires may add an important component to case-comparison studies; and in
some cases, using historical physical activity assessments may be more time- and cost-efficient than collecting physical activity data in a cohort and conducting a longitudinal study.

The lack of an acceptable criterion is a challenge in determining the validity of self-reported historical physical activity. A better approach may be to establish construct-related validity rather than criterion-related validity (18). Construct-related validity may be established by testing expected relations and predictable changes in variables to help determine overall accuracy. The purpose of this study was to establish construct-related validity for self-reported historical walking, running, and jogging (WRJ) activity engaged in up to 10 years prior to recall by using treadmill-determined cardiorespiratory fitness and by comparing data on previously measured health markers with recalled historical activity.

MATERIALS AND METHODS

Study design and participants

A research institute review board and a university institutional review board approved the protocol of this study. This analysis was based on data from the Aerobics Center Longitudinal Study. Participants were examined at a preventive medicine clinic in Dallas, Texas, at least once between 1976 and 1985 and then completed a mail-back follow-up questionnaire in 1986. The sample consisted of 4,100 men aged 18–80 years (mean = 47.5 years; standard deviation, 10.1) and 963 women aged 18–75 years (mean = 44.7 years; standard deviation, 10.4) at the time of initial medical examination.

Measurements

The medical examination included anthropometric measures such as body composition, height, and weight, as well as blood chemistry tests, blood pressure measurement, and a maximal exercise test. Body mass index was calculated by dividing weight in kilograms by squared height in meters. At the time of the medical examination, participants completed an extensive medical history questionnaire on personal and familial health. Questions assessing health habits (e.g., smoking, alcohol consumption, exercise) and general demographic information were included in the medical history questionnaire.

Cardiorespiratory fitness was measured by performance on a maximal treadmill test using a modified Balke protocol. Treadmill speed was set at 88 m/minute at a 0 percent grade for the first minute. Treadmill grade was increased to 2 percent for the second minute and increased by 1 percent for each minute thereafter for 25 minutes. After 25 minutes, treadmill grade was not changed, but speed was increased 5.4 m/minute for each minute until the end of the test. The test continued until the participant reached exhaustion or the supervising physician stopped the test for medical reasons. Subjects were encouraged to expend maximal effort. Participants unable to achieve at least 85 percent of their age-predicted maximal heart rate on the maximal exercise test were excluded from the analyses. Treadmill time from the clinic examination provided the measure of cardiorespiratory fitness for this study. Treadmill time with our protocol correlates highly ($r > 0.90$) with maximal oxygen uptake (19, 20); thus, we consider our measure of cardiorespiratory fitness analogous to aerobic power.

In 1986, a mail-back follow-up survey was sent to clinic patients who had responded to a previous mail-back survey in 1982. The purpose of the 1986 survey was to assess exercise patterns over the previous decade and orthopedic injuries. (See the Appendix for the WRJ recall question included in the mail-back survey.) Participants were asked whether they had engaged in WRJ in each year between 1976 and 1985. For each respective year, participants who responded that they had engaged in WRJ quantified the number of months in which they had participated in WRJ, the approximate number of workouts per week, the miles covered per workout, and the average mile time. Additionally, for each year between 1976 and 1985, participants were asked whether they had participated in strenuous sports other than WRJ (e.g., racquet sports, cycling, swimming, aerobic dancing, basketball, soccer, etc.) at least twice per week for 6 consecutive months. This information was used in the analysis to control for potential differences in treadmill test performance due to an effect of participation in strenuous activities other than WRJ.

Using the compendium of physical activities developed by Ainsworth et al. (21), we assigned metabolic equivalent (MET) values to participant responses based on the average mile time of reported WRJ. A MET is a unit of estimation of the energy cost of activity (22). MET values were multiplied by self-reported duration of WRJ (miles per workout multiplied by minutes per mile) and self-reported frequency (workouts per week and months per year) to convert self-reported historical physical activity to estimated MET-hours per week of historical WRJ averaged over the year for each year.

Data analysis

Participants may have visited the clinic more than once between 1976 and 1985. The only comparisons made were between variables measured during the clinic visit year and the specific, corresponding year of self-reported WRJ on the survey (e.g., if the person completed a treadmill test in 1978, these results were compared with the person’s WRJ reported in 1986 for 1978).

Statistical analyses used to obtain evidence of the construct-related validity of self-reported historical WRJ included calculation of Spearman correlations and partial correlations, analysis of variance, analysis of covariance, and $t$ tests. We calculated correlations to determine the relation between MET-hours per week of historical WRJ and treadmill time. For each year between 1976 and 1985, we used analysis of variance and analysis of covariance to compare the previously measured treadmill times of subjects classified as historically sufficiently active for a health benefit (engaging in ≥7.5 MET-hours/week) with the treadmill times of those classified as insufficiently active for a health benefit. Analyses were adjusted for age, body mass.
Validity of Recalled Historical Physical Activity

Am J Epidemiol 2004;160:279–286

INDEX, and participation in other strenuous activity. A cutoff level of 7.5 MET-hours per week was established as the minimum level of activity needed to be in compliance with Centers for Disease Control and Prevention/American College of Sports Medicine physical activity recommendations (23). A physical activity level of 7.5 MET-hours per week is approximately the energy cost of engaging in 30 minutes of moderate activity 5 days per week. Using this cutoff, participants were classified as historically either sufficiently or insufficiently physically active based on estimated MET-hours of WRJ per week.

Finally, we used t tests to investigate differences in health markers from the first clinic visit between participants classified as historically sufficiently active for a health benefit and those classified as insufficiently active. Health variables of interest were glucose, cholesterol, and triglyceride levels, resting systolic blood pressure, body mass index, and treadmill time. For each health variable, we conducted t tests by physical activity level for men and women separately. We conducted multiple t tests because not all participants had data on all health markers. Because of the number of correlations and t tests for which results are reported, we used a restricted alpha level for interpretation of statistical significance (i.e., p < 0.01).

RESULTS

More than 97 percent of subjects in this study were Caucasian. Baseline demographic and clinical characteristics of the participants as reported on their first visit to the clinic are shown in table 1. Average clinical measures for both men and women fell within normal ranges, with the exception of total cholesterol level. Total cholesterol level was borderline high for men (208 mmol/liter). On average, the treadmill test lasted 19.2 minutes for men and 13.5 minutes for women.

Correlation of historical WRJ with treadmill time

We calculated Spearman correlations to determine the association between MET-hours per week of historical WRJ and treadmill time. Correlations were calculated separately by sex for each year between 1976 and 1985. Sex-specific correlation coefficients for each year are presented in table 2. Moderate Spearman correlations relating treadmill time and MET-hours per week of historical WRJ were significant (p < 0.001 for all) for men and women for each of the 10 years. Results remained consistent after adjustment for age, body mass index, and participation in other

---

TABLE 1. Baseline characteristics of men and women at the first clinic visit, Aerobics Center Longitudinal Study, Dallas, Texas, 1976–1985*

<table>
<thead>
<tr>
<th>Measure</th>
<th>Men (n = 4,100)</th>
<th>Women (n = 963)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total no. of visits for 1976–1985</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1,721</td>
<td>598</td>
</tr>
<tr>
<td>2</td>
<td>758</td>
<td>150</td>
</tr>
<tr>
<td>3</td>
<td>443</td>
<td>81</td>
</tr>
<tr>
<td>4</td>
<td>328</td>
<td>51</td>
</tr>
<tr>
<td>≥5</td>
<td>850</td>
<td>83</td>
</tr>
<tr>
<td>Age (years)</td>
<td>47.5 (10.1)</td>
<td>44.7 (10.4)</td>
</tr>
<tr>
<td>Body mass index†</td>
<td>25.2 (3.0)</td>
<td>21.8 (3.1)</td>
</tr>
<tr>
<td>Total cholesterol level (mmol/liter)</td>
<td>208.3 (53.7)</td>
<td>199.2 (35.5)</td>
</tr>
<tr>
<td>High density lipoprotein cholesterol level (mmol/liter)</td>
<td>46.8 (13.8)</td>
<td>60.0 (12.9)</td>
</tr>
<tr>
<td>Glucose level (mg/dl)</td>
<td>99.1 (15.3)</td>
<td>92.9 (14.9)</td>
</tr>
<tr>
<td>Resting systolic blood pressure (mmHg)</td>
<td>119.8 (13.4)</td>
<td>111.6 (15.1)</td>
</tr>
<tr>
<td>Resting diastolic blood pressure (mmHg)</td>
<td>79.5 (8.9)</td>
<td>74.4 (9.3)</td>
</tr>
<tr>
<td>Treadmill time (minutes)</td>
<td>19.2 (5.0)</td>
<td>13.5 (4.6)</td>
</tr>
<tr>
<td>Walking, running, and jogging (metabolic equivalent-hours/ week)</td>
<td>12.9 (22.8)</td>
<td>8.0 (16.3)</td>
</tr>
</tbody>
</table>

* Measurements for some health markers (cholesterol level, glucose level, and resting systolic and diastolic blood pressures) were not obtained from some participants.
† SD, standard deviation.
‡ Weight (kg)/height (m)².
strenuous activity. For men, adjusted $r$'s ranged from 0.49 to 0.58, with a median value of 0.53. For women, adjusted $r$'s ranged from 0.40 to 0.61, with a median value of 0.47. The strength of the relation between treadmill time and self-reported historical WRJ did not appreciably diminish with a longer period of recall.

Self-report of sufficient levels of WRJ compared with insufficient WRJ

Analysis of variance was used to compare the mean treadmill time of men reporting insufficient historical WRJ with that of men reporting a sufficient level of WRJ for each year recalled. Mean treadmill times for men are presented in figure 1. There was a significant difference ($p < 0.001$) between the treadmill time of men reporting insufficient levels of WRJ and the treadmill time of men reporting sufficient WRJ for each year of recall. Results remained consistent after we conducted analysis of covariance and adjusted the data for age, body mass index, and participation in other strenuous activity. Men who reported sufficient WRJ had treadmill times 3.1–3.9 minutes longer than men who did not report participating in enough WRJ for a health benefit; the average effect size was 0.68 minutes (range, 0.56–0.75) across the 10-year recall period.

We also used analysis of variance to compare the mean treadmill time of women reporting insufficient historical WRJ with that of women reporting sufficient WRJ for each year. Mean treadmill times for women are presented in figure 2. There was a significant difference ($p < 0.001$) between the treadmill time of women reporting insufficient levels of historical WRJ and the treadmill time of women reporting levels of WRJ sufficient for a health benefit for each year of recall. Results remained consistent after we conducted analysis of covariance and adjusted for age, body mass index, and participation in other strenuous activity. Women who reported sufficient WRJ had treadmill times 2.7–4.5 minutes longer than women who reported insufficient WRJ; the average effect size was 0.81 minutes (range, 0.53–1.19) across the 10-year recall period.

![Figure 1](https://academic.oup.com/aje/article-abstract/160/3/279/59181)

**FIGURE 1.** Mean treadmill times of men reporting levels of historical walking, running, and jogging (WRJ) activity sufficient to obtain a health benefit (≥7.5 metabolic equivalent-hours/week) and men reporting insufficient levels of activity (<7.5 metabolic equivalent-hours/week), Aerobics Center Longitudinal Study, Dallas, Texas, 1976–1985. The difference between the groups was significant ($p < 0.0001$) for each year.
Relation between health markers and compliance with physical activity guidelines

On the basis of data from the first clinic visit between 1976 and 1985, we used $t$ tests to evaluate differences in measured health markers between persons classified as historically sufficiently physically active and those classified as insufficiently active. Results from the $t$ tests are presented in table 3. Men classified as engaging in sufficient levels of historical WRJ had significantly ($p \leq 0.001$) lower triglyceride levels and body mass indices and higher treadmill times than historically insufficiently active men. Women classified as historically sufficiently active had a significantly ($p < 0.001$) lower body mass index and a higher treadmill time.

DISCUSSION

Collectively, these results provide support for the construct-related validity of self-reported historical physical activity.

TABLE 3. Results of $t$ tests comparing mean levels of health markers at the first clinic visit for participants categorized as historically sufficiently active to receive a health benefit ($\geq 7.5$ MET*-hours/week of walking, running, and jogging activity) and participants characterized as historically insufficiently active ($<7.5$ MET-hours/week of walking, running, and jogging activity), Aerobics Center Longitudinal Study, Dallas, Texas, 1976–1985†

<table>
<thead>
<tr>
<th>Health marker</th>
<th>Men</th>
<th>Women</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Insufficiently active</td>
<td>Sufficiently active</td>
</tr>
<tr>
<td>Glucose level (mg/dl)</td>
<td>99.60 (17.76)§</td>
<td>103.86 (226.61)</td>
</tr>
<tr>
<td>Cholesterol level (mg/dl)</td>
<td>209.40 (37.40)</td>
<td>207.15 (67.20)</td>
</tr>
<tr>
<td>Triglyceride level (mg/dl)</td>
<td>139.31 (116.78)</td>
<td>110.93 (82.64)</td>
</tr>
<tr>
<td>Resting systolic blood pressure (mmHg)</td>
<td>119.45 (13.29)</td>
<td>120.12 (13.50)</td>
</tr>
<tr>
<td>Body mass index¶</td>
<td>25.86 (3.29)</td>
<td>24.51 (2.55)</td>
</tr>
<tr>
<td>Treadmill time (minutes)</td>
<td>16.86 (4.36)</td>
<td>21.82 (4.45)</td>
</tr>
</tbody>
</table>

* MET, metabolic equivalent.
† Information on some health markers was not obtained from some participants.
‡ Values were obtained by subtracting data for the sufficiently active from data for the insufficiently active.
§ Numbers in parentheses, standard deviation.
¶ Weight (kg)/height (m)$^2$. 

Am J Epidemiol 2004;160:279–286
activity. Historical WRJ may be validly assessed when participants are asked to quantify average time per mile of WRJ, average number of miles per workout, and average number of workouts per week over a period of time up to 10 years in the past. Correlations between MET-hours per week derived from self-reported historical WRJ for a particular year between 1976 and 1985 and treadmill time measured during the same year of recall were similar across years and did not appreciably diminish with a longer period of recall.

There are a variety of approaches for establishing construct-related validity. Investigating expected relations between variables and predictable changes in variables might provide convergent evidence for the validity of a behavior, trait, or quality that is not measurable by other methods of assessment. Correlation between self-reported historical physical activity and previously measured aerobic fitness contributes evidence of the validity of recalled historical WRJ. Participants reporting higher levels of historical WRJ, on average, had higher treadmill times than participants reporting lower levels of historical WRJ during the years for which physical activity was recalled. Participants classified as being historically sufficiently active for a health benefit had higher treadmill times and lower levels of certain health markers such as triglycerides and body mass index. With the exception of the health markers, results were similar between men and women. The lack of difference in health markers could be attributed to the small number of women in that analysis and their better health profiles at baseline.

Assessment of the validity of self-reported historical physical activity involves measurement of a criterion in subjects during the time period for which physical activity is to be recalled at some point later. In a study by Kohl et al. (24), responses to the same mail-back survey of historical physical activity as that used in the current study were compared with personal reports to a physician. The validation criterion in the Kohl et al. study was self-reported data on physical activity gathered during clinic visits from 1976 through 1985. Kohl et al. found that men consistently overreported historical miles per week and underreported average times per week. Results were inconsistent among women. The authors found that historical WRJ could be accurately recalled up to 7 years in the past when recalled activity was compared with personal reports to physicians.

A major strength of the current study is the uniqueness of the data. Rather than correlate and compare self-reported levels of historical physical activity with a subjective criterion, we assessed the validity of self-reported historical physical activity by use of an objective criterion measure. Information on treadmill time collected in the past was available for all subjects in the current study. While other investigators have compared recalled historical physical activity with physical activity self-reported (i.e., personally reported to a physician) at a prior point in time (25, 26), this study may have been among the first to compare self-reported historical physical activity with levels of aerobic fitness measured in the past.

We used fitness measured between 1976 and 1985 as the criterion for validation of self-reported historical WRJ, though there are some limitations involved in using fitness to validate self-reported physical activity. The association of physical activity with cardiorespiratory fitness is weakened by, among other factors, crude and imprecise measurement of physical activity and the genetic component of fitness (27). Low-to-moderate correlations between physical activity and physical fitness are to be expected. In the Surgeon General’s report on physical activity and health (1), a median correlation or concurrent validity coefficient of 0.41 was reported across 12 studies assessing the relations between measures of cardiorespiratory endurance and self-reported physical activity. The median correlations reported in the present study of 0.53 for males and 0.47 for females are higher, indicating moderate relations between cardiorespiratory endurance and self-reported physical activity, whether concurrently or historically recalled.

Quantifying historical physical activity is a difficult task, since physical activity behavior is multidimensional and is subject to several different standards of evaluation (28). Thus, the true relation between health status and historical physical activity remains an enigma to researchers, who are impeded by present measurement techniques (29). Durante and Ainsworth (30) stated that to obtain valid physical activity data via self-report, it is important to assess the types of activities engaged in, the frequency of the activities, and the time period in which respondents participated in the activities. While surveys eliciting recall of physical activity engaged in over a relatively brief period of time may quantify the frequency and duration of activity with a benefit of less vulnerability to recall bias (18), a recall instrument with a longer historical time frame might assess regular patterns of exposure to physical activity (31).

Errors in recall of physical activity can be made for a number of reasons. Estimation of historical physical activity can be affected by cognitive factors such as a person’s ability to store and retrieve information (32, 33). Individuals may have differing interpretations of self-report questions. Some self-report items depend on respondents’ being familiar with definitions of activity and exercise, as well as their being able to distinguish between different exercise intensities (34). Respondents to physical activity recall questions may tend to overestimate aerobic activity and underestimate sedentary activity (29). However, despite the factors limiting recall of prior physical activity, assessment of historical physical activity by self-report can be done with reasonable precision and accuracy (15, 24, 35).

While the validity of self-reported historical physical activity, not its reliability, was assessed, it is well known that for a measurement to be valid it must also possess some degree of reliability. Therefore, the reliability of self-reported historical physical activity may be ascertained from the findings of prior research. Other studies have found moderate-to-high test-retest reliability for multiple lifetime physical activity questionnaires when the questionnaires were administered two times 3 months–1 year apart. Self-report historical physical activity questionnaires have been found to have high reproducibility when different types of activity, time periods, and intensity levels are being assessed (15, 35, 36).

The representativeness of this sample should be a consideration in application of the results of this study. The particip-
participants were mostly White males with higher levels of education and socioeconomic status. People with a higher educational level and socioeconomic status may report historical physical activity more accurately, thus limiting the generalizability of these results to populations with lower levels of education or lower socioeconomic status. Future research should be designed to establish the validity of self-reported historical physical activity for populations of differing educational levels, socioeconomic status levels, and ethnicities.

In summary, historical physical activity may be an important factor in chronic disease research. The results of this study were consistent for men and women, with little change being seen after adjustment for age, body mass index, and participation in other strenuous physical activity. It was found in this study that historical WRJ can be assessed with reasonable validity when compared with treadmill performance, and there is no decay in the accuracy of this reporting for up to 10 years.

ACKNOWLEDGMENTS

This study was supported in part by National Institutes of Health grants AG06945 and HL-47478.

The authors thank Beth Wright at the Cooper Institute for her assistance with manuscript preparation; Dr. H. W. Kohl of the Centers for Disease Control and Prevention for his important role in the development and administration of the 1986 mail-back survey; and Dr. Marilu Meredith and the data entry staff of the Cooper Institute for data management.

REFERENCES


APPENDIX

Questions on Historical Physical Activity

Please give us your run, walk, or jog participation from 1976–1985. For each year, circle YES if you participated in a run, walk, or jog exercise program and NO if you did not participate. If you circle YES, please give the approximate number of months per year, workouts per week, miles covered per workout, and the average time per mile. Please provide a complete answer for each year.

<table>
<thead>
<tr>
<th>Year</th>
<th>Walk, run, or jog?</th>
<th>No. of months per year</th>
<th>No. of workouts per week</th>
<th>Average miles per workout</th>
<th>Average minutes per mile</th>
</tr>
</thead>
<tbody>
<tr>
<td>1976</td>
<td>NO YES</td>
<td>→</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1977</td>
<td>NO YES</td>
<td>→</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1978</td>
<td>NO YES</td>
<td>→</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1979</td>
<td>NO YES</td>
<td>→</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1980</td>
<td>NO YES</td>
<td>→</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1981</td>
<td>NO YES</td>
<td>→</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1982</td>
<td>NO YES</td>
<td>→</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1983</td>
<td>NO YES</td>
<td>→</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1984</td>
<td>NO YES</td>
<td>→</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1985</td>
<td>NO YES</td>
<td>→</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Have you ever participated at least twice a week for 6 consecutive months in strenuous sports other than running, walking, and jogging? (Examples: racquet sports, cycling, swimming, aerobic dance, or strenuous sports involving running such as basketball, soccer, etc.) Please circle your answer. If you circle YES, please give the year you began your program and indicate the activity you did most often.

YES→ What year did you begin this program? 19____
What activity did you do most often? ________________________________________________________________

NO

In what years between 1976 and 1985 did you participate at least twice a week for 6 consecutive months in strenuous sports other than running, walking, and jogging? Please circle all years that apply.