A Randomized Walking Trial in Postmenopausal Women

Effects on Physical Activity and Health 10 Years Later

Mark A. Pereira, PhD; Andrea M. Kriska, PhD; Richard D. Day, PhD; Jane A. Cauley, DrPH; Ronald E. LaPorte, PhD; Lewis H. Kuller, DrPH, MD

Background: It is important to determine if permanent lifestyle changes may result from physical activity interventions and whether health may be affected by these changes.

Objective: To conduct a 10-year follow-up of physical activity and self-reported health status in participants of a randomized clinical trial of walking intervention.

Methods: Of the original 229 volunteer postmenopausal women who participated in the original clinical trial, 196 (N = 96 intervention and 100 controls) completed the 10-year follow-up telephone interview. The interview protocol included questions on self-reported walking for exercise and purposes other than exercise, the Paffenbarger sport and exercise index, functional status, and various chronic diseases and conditions.

Results: The median values for both usual walking for exercise and total walking were significantly higher for walkers compared with controls (for both, $P = .01$), with median differences of 706 and 420 kcal/wk, respectively. After excluding women who reported heart disease during the original trial, 2 women in the walking group (2%) and 11 women in the control group (12%) reported physician-diagnosed heart disease over the last 10 years ($P = .07$). There were also fewer hospitalizations, surgeries, and falls among women in the walking group, although these differences were not statistically significant ($P > .05$).

Conclusions: Although limited by self-report, this study may be the first to demonstrate long-term exercise compliance to a randomized control trial in older women and to suggest that health benefits may have ensued as a result of these increased activity levels.

Arch Intern Med. 1998;158:1695-1701

The First Physical Activity and Health Report of the Surgeon General has emphasized the importance of a physically active lifestyle in the prevention of many noncommunicable diseases and their risk factors.1 Despite the well-documented importance of physical activity for general health and well-being, a large portion of US citizens lead sedentary lives. Physical activity decreases with age, and older women are typically the least active.2-5

There have been few long-term (≥1 year) randomized exercise trials in older women who demonstrated adequate compliance to the intervention.6-9 In fact, most programs have not exceeded a duration of 6 months, and even fewer have gathered follow-up data pertaining to the physical activity habits and health status of the program participants beyond the cessation of the intervention. Although there have been some follow-up studies of randomized exercise trials,10-11 we are unaware of a long-term follow-up study to a previously successful randomized exercise trial in older women. From a public health perspective it is relevant to determine if participants are likely to make permanent lifestyle changes as a result of a successful physical activity intervention. It is also important to determine if indexes of health status have been favorably affected by these lifestyle changes.

The primary aim of our study was to test the hypothesis that the women originally randomized to the walking group of a clinical trial from 1982 to 1985 were still walking more 10 years after the closure of the trial than the women originally randomized to the control group. Exploratory analyses were also pursued to examine types of physical activities other than walking, self-reported physician-diagnosed diseases and conditions, and other health outcome parameters between walkers and controls.

RESULTS

Sample Contact and Interviews

Nine of the original 229 randomized women were confirmed to be dead (4 from the walking group and 5 from the control group).
METHODS

This study is a 10-year follow-up to a randomized clinical trial of walking intervention in postmenopausal women. The description of the original trial follows.

OVERVIEW OF ORIGINAL CLINICAL TRIAL

Fifteen years ago a randomized control trial was initiated (1982-1985) to examine the effects of walking activity on bone parameters in postmenopausal women. Two hundred twenty-nine apparently healthy, predominantly white, middle-class volunteers residing in the Pittsburgh, Pa, area were randomized to either the walking intervention (n = 114) or control (n = 115) group. For entry into the study, the women had to meet 4 primary eligibility criteria: aged 50 to 65 years; at least 1 year after cessation of menses; abstention from estrogen replacement therapy; and freedom from limitations that might preclude walking.

An important result of this randomized walking trial was the excellent compliance to the exercise intervention as reported on monthly walking logs. Eighty percent of the intervention women averaged 5 or more miles/wk over the first 2 years of the trial, with 61% averaging 7 or more miles/wk. The success of the intervention was validated by independent activity measures, including subjective estimates obtained from the Paffenbarger Physical Activity Questionnaire and objective measures determined by the Large Scale Integrated Activity Monitor.

1995 TELEPHONE INTERVIEW FOLLOW-UP STUDY

Our study focused on locating and recontacting the participants of the original clinical trial with the purpose of obtaining physical activity and health status data using a comprehensive telephone interview–administered questionnaire. Stored records and databases from the original trial were used to collect descriptive information and data from the original 229 women who were randomized at baseline of the original clinical trial (1982-1983). Prior to any attempts to recontact the women, a nationwide death search was used to identify any women who had died.

Two middle-aged women assisted in both pilot testing the questionnaire and contacting and interviewing the participants. These telephone interviewers remained masked to the original group assignment of the participants until the end of each interview. Our study was approved by the Biomedical Institutional Review Board at the University of Pittsburgh, Pittsburgh, and all participants provided verbal consent to complete the telephone interviews.

The telephone interview consisted of 4 general sections: (1) physical activity, (2) functional ability, (3) health status, and (4) behavioral and descriptive characteristics. Most items in the interview were adapted either from existing questionnaires and health surveys used in the original study or from questionnaires used elsewhere in epidemiological investigations of older women and national surveys of the United States.

Physical Activity

The Paffenbarger Physical Activity Questionnaire was previously used to assess physical activity levels in participants of the original clinical trial. The reliability and validity of this questionnaire has been demonstrated in this group of postmenopausal women. The only modification that was made to the version of the Paffenbarger questionnaire administered during the original trial was the question inquiring about the weekly number of city blocks walked. A modification of this question from the Study of Osteoporotic Fractures was used. The modification partitions the walking question into 2 parts, inquiring about the number of blocks walked for exercise (“Do you take walks for exercise?” If yes: “On an average day, how many minutes during that day do you walk for exercise?”) and the number of blocks walked apart from exercise for purposes such as going to the store, church, or walking a dog (“In addition to walks for exercise, on an average day, how many minutes do you walk to the store, to church, to a friend’s house, or work, including walking your dog?”). These 2 walking estimates—“usual walking for exercise” and “non–exercise walking”—were analyzed independently as well as collapsed into a total walking estimate (total usual walking). Each 20 minutes of walking was assumed to...

CURRENT PHYSICAL ACTIVITY LEVELS OF WALKERS AND CONTROLS

Walking for Exercise

Based on the 4 different ways in which walking was quantified, the women originally randomized to the walking intervention group consistently reported more walking in 1995 than those randomized to the control group. As shown in Figure 1, the estimate of weekly kilocalorie expenditure for total usual walking revealed a significantly higher median and 75th percentile for the walkers than for the controls (median, 1344 vs 924 kcal/wk; \( P = .01 \)). The results for usual walking for exercise revealed an even larger median difference (1008 vs 302 kcal/wk; \( P = .01 \)). Although the walking estimate from the sports and recreation index of the Paffenbarger questionnaire also demonstrated higher values for walkers,
equal 1 mile or 96 kcal. The amount of walking for a typical day was then multiplied by 7 to derive a weekly estimate (kilocalorie per week). Sports and recreational activities were weighted by an estimate of their relative metabolic cost. The sport section of the Paffenbarger questionnaire also included walking, which was also analyzed as a separate variable in addition to its contribution to the sport index.

Additional questions were developed during pilot testing to inquire specifically about walking activity. These questions included the length of time the individual could walk before having to stop, a self-comparison of present levels of walking to the amount performed 10 years ago, and the frequency and duration of walking for exercise over the last year. This latter question differed from usual walking for exercise in that it inquired about minutes of walking on a typical day (“Have you walked for exercise over the past year?” If yes: “How often do you usually walk for exercise in a week? Times per week? Minutes per walk?”). As in usual walking for exercise, 20 minutes was considered to be the equivalent of 1 mile (96 kcal).

Functional Ability

Questions on functional ability and periods of confinement inquired about the level of difficulty that the individual had in completing various daily tasks, including whether the individual needed the help of a cane, walker, or another individual to walk or climb stairs. We also asked about the length of time the individual could walk at a steady pace without becoming tired and having to stop.

Self-rated Health

Participants were asked to rate their overall health relative to other women their own age according to the following question: “In comparison to other women your age, how would you rate your overall health?”

The question “Has a doctor ever told you that you had (condition or disease)?” was preceded by the phrase “over the past 10 years” to differentiate diseases or conditions that were present before or during the original clinical trial from those that developed since the closure of the trial. If the individual answered yes to any of these questions the interviewer proceeded to inquire about the year of diagnosis and treatment for that disease or condition. The baseline records were reviewed in an attempt to eliminate those women who had the disease or condition before the start of the randomized clinical trial. Questions regarding hospitalizations, surgeries, falls, and medications over the last 10 years were also asked.

Other Behavioral and Descriptive Characteristics

Questions on cigarette smoking and alcohol consumption were adapted from the questions used during the original trial. Questions on current height, weight, employment or volunteer work, and whether the individual was living alone were asked at the end of the interview.

STATISTICAL ANALYSIS

Data were analyzed using unpaired t-tests to compare means of continuous variables that were distributed gaussian. The Wilcoxon rank sum test for 2 group comparisons and the Kruskal-Wallis test for comparing more than 2 groups with the Dunn procedure for subsequent pairwise comparisons were used for analysis of kilocalorie scores because of their skewed distributions. The $\chi^2$ test was used to compare percentages between groups for data distributed binomially or ordinally. Ten-year cumulative incidence rates were computed for reported diseases and conditions, along with relative risks and 95% confidence intervals. A Bonferroni correction of $P$ values was used to maintain an experimentwise type I error rate of approximately 0.05 according to the following formula: corrected $P = 1 - (1 - p)^n$, where $p$ is the original $P$ value and $n$ is the number of comparisons being made. To avoid inflation of the type II error rate, the Bonferroni correction was computed specific to each construct in which more than 1 comparison was being made. For example, when the 4 estimates of walking were being examined, the expression to correct the $P$ value for each measure was $1 - (1 - p)^4$, with corrected $P < 0.05$ deemed statistically significant. The SAS statistical software was used for all analyses.
steps appeared to be different, with 7% of walkers and 13% of controls reporting any difficulty. Fifty-four percent of walkers and 42% of controls rated their overall health as “very good.” None of the differences between walkers and controls were found to be statistically significant ($P > .05$).

### Diseases and Other Conditions

As shown in Table 3, 2.0% of the women in the walking group reported heart disease compared with 12.0% of the women in the control group since the end of the original trial. This apparent difference, although not statistically significant after the Bonferroni correction ($P = .07$), did not appear to be explained by hormone use or cigarette smoking. Reported use of estrogen and progesterone was not significantly different between walkers and controls (Table 1). Only 1 of 11 women with self-reported heart disease in the control group was a current smoker at baseline of the original trial and 3 were classified as ever-smokers at the 10-year follow-up. The difference between groups in reported angina was consistent with the difference between groups in reported heart disease.

Twenty-five percent of the women in the walking group and 37% of the women in the control group reported 1 or more hospitalizations over the last 10 years. Orthopedic surgeries were reported by 6.3% of walkers and 16.0% of controls. Consistent with the self-reported heart disease and angina, 3.0% of controls reported heart surgery, while none of the walkers reported heart surgery. The occurrence of falls over the last 12 months was 27.1% in walkers and 33.0% in controls, while the percentage of women reporting more than 1 fall over the last 12 months was 23.1% in walkers and 30.3% in controls. None of these differences were statistically significant ($P > .05$).

### TRACKING OF PHYSICAL ACTIVITY FROM BASELINE TO FOLLOW-UP

As identified from the original clinical trial, walking at baseline was positively associated with continued walking over the course of the intervention study. Due to the predictive value of activity at baseline for compliance to the in-
intervention, we were interested in examining reported levels of walking at follow-up stratified by baseline walking status in the control group and compliance to the intervention in the walking group. Reported blocks walked per day at baseline of the original intervention were used to classify the women in the control group as “inactive” or “active.” Control women walking less than or equal to 5.1 blocks per day were classified as inactive at baseline, while controls walking more than 5.1 blocks per day were classified as active at baseline. This cutoff point resulted in a similar number of women in the control group classified as active at baseline. This cutoff point resulted in a similar number of women in the control group classified as active (n = 55) relative to “compliers” in the walking group. Reported blocks walked per day at baseline and end of trial and usual walking for exercise reported during the follow-up. Asterisk indicates P < .05 compared with all other groups.

The differences in reported walking between the 2 randomized groups across the various walking measures provide support for a long-term effect of the original intervention on the current physical activity levels of these women. Those women who comprised the walking intervention group were more active than the control women at the end of the trial and maintained relatively higher levels of activity a decade later. To our knowledge, this is the first study that has demonstrated long-term exercise compliance in an older population 10 years after the intervention effort ended, suggesting that a permanent lifestyle change had been made.

Because walking was the intervention mode used in the original clinical trial and it is the most popular leisure-time physical activity in older adults,29-31 it is not surprising that the greatest difference in activity between the randomized groups was found for walking activity. The popularity of walking in older adults is most likely a reflection of the feasibility of walking as an activity that can be incorporated into the daily routine. Furthermore, in

Table 2. Comparison of Functional Ability Variables Between Walkers and Controls 10 Years After the Closure of the Original Clinical Trial

<table>
<thead>
<tr>
<th>Variable</th>
<th>Walkers (n = 96)</th>
<th>Controls (n = 100)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Duration of steady walking, min*</td>
<td>45.0 (30.0, 60.0)</td>
<td>45.0 (25.0, 60.0)</td>
<td>.99</td>
</tr>
<tr>
<td>Uses cane, walker, or other aid,%</td>
<td>4.2</td>
<td>7.0</td>
<td>.95</td>
</tr>
<tr>
<td>With difficulty, %</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Walking 2 or 3 blocks</td>
<td>9.4</td>
<td>16.0</td>
<td>.65</td>
</tr>
<tr>
<td>Climbing up 10 steps</td>
<td>15.6</td>
<td>13.0</td>
<td>.99</td>
</tr>
<tr>
<td>Climbing down 10 steps</td>
<td>7.3</td>
<td>13.0</td>
<td>.72</td>
</tr>
<tr>
<td>Doing heavy housework</td>
<td>21.9</td>
<td>25.0</td>
<td>.99</td>
</tr>
</tbody>
</table>

* Two or 3 blocks outside on level ground; median (25th, 75th percentile). 94 walkers and 99 controls.

Table 3. Comparison of Self-reported Chronic Diseases or Conditions Diagnosed by a Physician Over the Last 10 Years in Walkers and Controls

<table>
<thead>
<tr>
<th>Disease or Condition</th>
<th>Walkers</th>
<th>Controls</th>
<th>RR (95% CI)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heart disease</td>
<td>2/93 = 0.02</td>
<td>11/93 = 0.12</td>
<td>0.18 (0.04-0.80)</td>
</tr>
<tr>
<td>Chest pain or angina</td>
<td>2/94 = 0.02</td>
<td>8/95 = 0.08</td>
<td>0.25 (0.06-1.16)</td>
</tr>
<tr>
<td>High blood pressure</td>
<td>14/82 = 0.17</td>
<td>15/79 = 0.19</td>
<td>0.90 (0.47-1.74)</td>
</tr>
<tr>
<td>Stroke</td>
<td>3/93 = 0.03</td>
<td>2/98 = 0.02</td>
<td>1.56 (0.27-9.15)</td>
</tr>
</tbody>
</table>

* RR indicates relative risk (incidence in walkers divided by incidence in controls); CI, confidence interval.

Comment

The differences in reported walking between the 2 randomized groups across the various walking measures provide support for a long-term effect of the original intervention on the current physical activity levels of these women. Those women who comprised the walking intervention group were more active than the control women at the end of the trial and maintained relatively higher levels of activity a decade later. To our knowledge, this is the first study that has demonstrated long-term exercise compliance in an older population 10 years after the intervention effort ended, suggesting that a permanent lifestyle change had been made.

Because walking was the intervention mode used in the original clinical trial and it is the most popular leisure-time physical activity in older adults,29-31 it is not surprising that the greatest difference in activity between the randomized groups was found for walking activity. The popularity of walking in older adults is most likely a reflection of the feasibility of walking as an activity that can be incorporated into the daily routine. Furthermore, in
comparison with other modes of exercise, such as jogging, walking has a relatively low risk of injury. In essence, the characteristics that describe walking—popularity, feasibility, and safety—are factors that have been found to be strongly associated with exercise adoption and compliance and are likely to explain some of the success of the exercise intervention.

Of the 10 most common activities reported for the Paffenbarger past-year sport and recreation index, the prevalence of 7 activities was higher in the walkers than the controls. Although the differences found in the Paffenbarger sports and recreation index were not statistically significant, walkers demonstrated higher medians and 75th percentiles compared with controls. Furthermore, the difference between groups in the Paffenbarger past-year sports and recreation kilocalorie score was still demonstrated after the exclusion of walking for exercise. These apparent differences between walkers and controls in activities other than walking are suggestive of an effect of the original intervention on general physical activity levels. Although the intervention focused solely on walking, there may have been a ripple effect, in that women in the walking group appear to have been participating in more sports and recreational activities than those in the control group.

An observation was made during the original clinical trial (A.M.K., unpublished data) that one effect of the intervention may have been the ability to recall activity more accurately. There were several cases in which the control women reported walking several miles in an unrealistic period of time, with no similar cases noted in the intervention women. This observation was later supported by data from Cauley et al, who found statistically significant correlations between blocks walked per day from the Paffenbarger questionnaire and the Large Scale Integrated activity monitor in the walking group, but not in the control group of the original clinical trial. Given these findings, it is possible that the difference in current levels of reported walking between the 2 groups is actually larger in magnitude than the data of our study indicate. However, further follow-up of this cohort with objective physical activity monitoring is required to adequately address this issue.

The results of the exploratory analyses examining the incidence of diseases and conditions in this study were limited by the capacity of the participants to accurately report the type of condition and time of diagnosis, the small sample size, and the lack of validation by medical records. These limitations warrant an additional follow-up study including objective clinical measures of physiological parameters and cardiovascular disease risk factors. Still, the finding for heart disease is consistent with the wealth of literature demonstrating the effect of moderate physical activity, such as walking, on the risk of cardiovascular disease. Although there was no effect from the intervention of the original clinical trial on blood lipid levels, the current findings of heart disease could be explained by other mechanisms through which moderate physical activity may reduce the risk for heart disease. These mechanisms include alterations in blood pressure, hemostasis, and insulin and glucose metabolism that could have affected the progression of atherosclerosis over the last 10 years, as well as direct effects on the myocardium that could have increased vascularization and/or attenuated sympathetic activity. It is also possible that the difference in physical activity between groups, which appears to be approximately 700 kcal of walking per week (~140 minutes or ~7 miles), maintained over a decade has resulted in lipoprotein differences that were not apparent at the conclusion of the original trial after 2.5 years of intervention.

Posner et al administered a 24-month cycling intervention in their exercise trial of older men and women. The length of the intervention and strategies to promote compliance in this study were similar to those of the original clinical trial of the present follow-up study. Although they did not report on follow-up data, those authors found a difference in the incidence of physician-diagnosed cardiac conditions between the exercise and control groups of 2.4% and 12.9%, respectively (P < .05).

In addition to heart disease, there were other patterns (although not statistically significant) in our data suggesting a difference in health status between walkers and controls in this follow-up study. Differences were found for hospitalizations, orthopedic surgeries, and miscellaneous surgeries reported over the last 10 years, in which the controls reported up to twice as many as the walkers. These general findings could be real differences between groups resulting from musculoskeletal conditioning from moderate physical activity over the 13-year period (total duration of trial and follow-up interval), as well as a more generalized effect of activity on physiological changes associated with aging.

One of the most impressive findings from this study was that the intervention was able to increase the activity levels of the least active women from the intervention group to the level of the most active women from the control group. Thus, the successful walking intervention not only increased the activity levels of those who were already relatively active but also appears to have elevated the activity levels of initially inactive participants to a level comparable with active women not exposed to the intervention.

Although no statistically significant differences were found between walkers and controls in self-reported functional ability, the numbers were small and this cohort may still be younger than the age when functional ability declines to the point that difficulties become prevalent enough to detect significant differences. There is good evidence to suggest that physical activity is a predictor of future functional status even when functional status and comorbidities are controlled for at baseline.

A major strength of this study was that 86% of the original participants were successfully recontacted and interviewed a decade later. The women in our study were relatively educated, healthy, postmenopausal white women who were residing in the Pittsburgh area when the original clinical trial took place from 1982-1985. Limitations of our study include the reliance on self-reported data, the use of a telephone interview to collect data, and the relatively small number of women from which to draw conclusions about disease end points.

Based on the success and feasibility of the exercise intervention of the original clinical trial and the high response rate to the 10-year follow-up telephone survey, this study may be the only investigation of its kind in any popu-
lation to date. The findings suggest that a permanent lifestyle change had been made as a result of a randomized clinical trial in older women and that health benefits may have ensued as a result of the increased activity levels.

Accepted for publication January 7, 1998.

This study was supported by a grant from the Sandoz Foundation for Gerontological Research, East Hanover, NJ. We acknowledge the hard work of Amy Rudzik and Sheree Shaffer in recontacting and interviewing the study participants. The work of Rivka Sandler, PhD, Jean McCormich, and Regina Onda, PhD, during the trial and/or in locating the data from the original trial is greatly appreciated. Comments on the manuscript by M. Kathryn H. Schmitz, PhD, are greatly appreciated.

Corresponding author: Mark A. Pereira, PhD, Division of Epidemiology, University of Minnesota, Suite 300, 1300 S Second St, Minneapolis, MN 55454 (e-mail: pereira@epivax.umn.edu).

REFERENCES