Reduced Disability and Mortality Among Aging Runners

A 21-Year Longitudinal Study

Eliza F. Chakravarty, MD, MS; Helen B. Hubert, PhD; Vijaya B. Lingala, PhD; James F. Fries, MD

Background: Exercise has been shown to improve many health outcomes and well-being of people of all ages. Long-term studies in older adults are needed to confirm disability and survival benefits of exercise.

Methods: Annual self-administered questionnaires were sent to 538 members of a nationwide running club and 423 healthy controls from northern California who were 50 years and older beginning in 1984. Data included running and exercise frequency, body mass index, and disability assessed by the Health Assessment Questionnaire Disability Index (HAQ-DI; scored from 0 [no difficulty] to 3 [unable to perform]) through 2005. A total of 284 runners and 156 controls completed the 21-year follow-up. Causes of death through 2003 were ascertained using the National Death Index. Multivariate regression techniques compared groups on disability and mortality.

Results: At baseline, runners were younger, leaner, and less likely to smoke compared with controls. The mean (SD) HAQ-DI score was higher for controls than for runners at all time points and increased with age in both groups, but to a lesser degree in runners (0.17 [0.34]) than in controls (0.36 [0.55]) (P < .001). Multivariate analyses showed that runners had a significantly lower risk of an HAQ-DI score of 0.5 (hazard ratio, 0.62; 95% confidence interval, 0.46-0.84). At 19 years, 15% of runners had died compared with 34% of controls. After adjustment for covariates, runners demonstrated a survival benefit (hazard ratio, 0.61; 95% confidence interval, 0.45-0.82). Disability and survival curves continued to diverge between groups after the 21-year follow-up as participants approached their ninth decade of life.

Conclusion: Vigorous exercise (running) at middle and older ages is associated with reduced disability in later life and a notable survival advantage.

Arch Intern Med. 2008;168(15):1638-1646

Author Affiliations: Division of Immunology and Rheumatology, Department of Medicine, Stanford University School of Medicine, Stanford, California.

Methods: Subjects were recruited in January 1984 to participate in a longitudinal study of the effect of long-distance running on health outcomes. Runners 50 years and older were enrolled from a nationwide running club, the 50+ Runners Association. Control subjects were recruited from the roster of the Stanford University Lipid Research Clinic Prevalence Study (LRC), which identified all permanent university staff and faculty aged between 26 and 70 years. This group was chosen to provide a sample of subjects with demographic characteristics similar to those of the running club. Details regarding the development of the cohort have been described elsewhere. Briefly, study descriptors were sent to all 1311 runners club and 2181 Stanford LRC participants in January 1984. Of these, 694 members of the runners club (run-
controls, performed at the baseline visit, showed excellent correlation between self-reported data and that obtained by physicians or trained observers. The LRC subjects who reported regular vigorous exercise, including running, were retained as controls so that the group represented an average of exercise habits in the community. Baseline variables and disability of subjects who withdrew from the study were compared with those who completed the 21-year follow-up. All subjects provided informed consent prior to participation.

GROUP ASSIGNMENT BY RUNNING HISTORY

The primary longitudinal analysis focused on the 284 runners and 156 controls who continued to participate in 2005. However, to control for potential self-selection bias based on group membership rather than running status and to include those who began to run but later discontinued, we created groups of “ever runners” and “never runners” based on the following question, “Have you ever run for exercise for a period of greater than 1 month?” Here, the ever-runners group includes subjects who currently run regularly but are not necessarily members of the runners club or those who may have run in the past but discontinued before study onset. This procedure shifted 143 participants (73 participating in 2005) from the control group to the ever-runners group. Over time, all groups decreased running activity, but the runners groups continued to accumulate more minutes per week of vigorous activity of all kinds.

ASSESSMENT OF OUTCOMES

Disability

The HAQ-DI is a self-reported instrument assessing functional ability in 8 areas: rising, dressing and grooming, hygiene, eating, walking, reach, grip, and routine physical activities. Each area is scored from 0 (no difficulty) to 3 (unable to perform), accounting for the use of special aids or devices or the assistance of another person. The HAQ-DI score is the mean scores of the 8 areas. It has been validated in numerous studies, is sensitive to change, and is widely used in observational studies and clinical trials.

Mortality

Mortality data were obtained for 100% of original participants from annual searches of the National Death Index through 2003. Data on death were ascertained even if participants had previously withdrawn from the study. The principal cause of death was determined using the National Death Index Plus service.

Statistical Analyses

Analyses were performed using SAS version 9.1 statistical software (SAS Institute Inc, Cary, North Carolina). Institutional review board approval was obtained before initiating the study. Characteristics of study participants by group membership (runners vs controls) are listed in Table 1. Values for all subjects at study inception are listed in the first columns. Baseline values (1984) for those subjects in each group who continued to participate through 2005 (completers) and for those who did not complete the follow-up period (noncompleters) are listed in subsequent columns. The results are given as means, with their standard errors, and 95% confidence intervals. Comparisons were considered statistically significant at P < .05.
similar between the 2 groups. Both groups had little disability at baseline, but runners had lower HAQ-DI scores and were more likely to have an HAQ-DI score of 0. Analysis of completers and noncompleters showed that, among controls, completers tended to be younger (P<.001), run more (P<.001), have less baseline disability (P<.05), and were more likely to have a baseline HAQ-DI score of 0 compared with noncompleters. Among runners, the only statistically significant difference was the 2-year age disparity between completers and noncompleters (P<.001).

Similar findings were observed when participants were divided into ever-runner and never-runner groups [Table 2].

Figure 1 shows the progression of mean disability levels by year. Mean disability levels increased with time in all groups. Figure 1A shows mean annual disability levels for runners compared with controls; Figure 1B il-
illustrates disability curves for ever-runner vs never-runner groups. Separate curves are shown for means computed using data available for all 961 initial study participants and for means computed for only those 440 subjects who completed the 21-year follow-up.

Members of the running groups had significantly lower mean disability levels at all time points. Members of both running groups had nearly identical mean disability levels irrespective of completer status, indicating few differences in disability between completers and those who dropped out or died.

In contrast, there were significant differences in disability levels between the inception cohort and completers among the control groups in both Figure 1A and B. Baseline disability levels in 1984 were statistically lower when computed for completers than when all initial study participants were included. This disability pattern among completers in the control groups continued at almost all time points, indicat-

![Figure 1. Mean disability levels by year. Solid lines represent data for all initial participants, and dashed lines represent data for those who continued participation through 2005. A, Runners club members vs community control subjects. B, Ever runners vs never runners. Error bars indicate SD.](image-url)
ing differential dropout of the subjects with higher disabil-
ity among the control groups, creating a potential bias to-
ward lower disability in the observed control groups. How-
ever, even when restricting the cohort to completers, runners 
had significantly lower disability compared with controls, 
and disability curves continued to diverge at 21 years of follow-
up. Analyses by ever- vs never-runners showed comparable 
results.

Mean disability levels for completers in each group (run-
ners and controls) are separated by sex in Figure 2. Both 
male and female runners maintained low disability levels 
at all time points, which were significantly lower than those of controls. The difference between runners and controls 
was most striking for women. Male controls had higher disabil-
ity levels than male runners at all time points except during the initial few years of the study. Few differences existed between male and female runners.

The rate of disability progression over 21 years of ob-
servation using general linear mixed models is shown in 
Figure 3. The rate of disability progression was signifi-

Figure 2. Mean disability levels by year separated by sex. Solid lines represent data for runners, and dashed lines represent data for controls who continued participation through 2005. Only subjects who completed the 21-year follow-up are included. Error bars indicate SD.

Figure 3. Progression of disability (PD). Linear mixed models of PD and postponement of disability. Regression lines are derived from linear mixed models and adjusted for the following covariates: age, sex, body mass index, smoking, and initial disability level. The PD is defined as the absolute difference between the 2 groups in the time required to cross a given level of disability. The example shown is to reach a Health Assessment Questionnaire Disability Index score of 0.15.
cantly lower for runners (0.007 points per year) compared with controls (0.016 points per year) (P < .001).

The time required to reach specified levels of disability was significantly longer for runners than for controls. The mean time to reaching an HAQ-DI score of 0.075 from study onset was approximately 2.6 years for controls and 8.7 years for runners, yielding a difference of approximately 6.2 years (95% CI, 3.9-8.9 years). Similarly, the time to reach an HAQ-DI score of 0.10 was 8.2 years (95% CI, 5.1-11.7 years) and to reach an HAQ-DI score of 0.15 was 12.1 years (95% CI, 8.1-18.3 years) for runners. These data illustrate that the slower rate of disability progression continued to increase over time among runners through at least the 21 years of observation.

Results of multivariate Cox regression analyses using time-dependent covariates are given in Table 3. The final model for disability outcomes (HAQ-DI scores of 0.5 and 1.0) included the following variables: group membership, age (year), sex, BMI (lagged by 1 year), and weekly vigorous exercise minutes (lagged by 1 year). These analyses were restricted to participants (completers and noncompleters) with a baseline HAQ-DI score of 0. Variables included in the final model include group (runners vs controls), age (year), sex, BMI (measured 1 year prior to disability measure), and vigorous exercise (measured 1 year prior to disability measure).

Analysis includes all participants at study inception. The final model includes group, age (year), sex, and baseline HAQ-DI score of 0.1. Smoking, BMI, and exercise did not meet statistical significance to be included in the final model.

Table 3. Multivariate Cox Regression Analyses for Disability and Mortality

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard Ratio (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HAQ-DI Score of 0.5&lt;sup&gt;a&lt;/sup&gt;</td>
</tr>
<tr>
<td>Runner</td>
<td>0.62 (0.46-0.84)</td>
</tr>
<tr>
<td>Age, y</td>
<td>1.07 (1.05-1.09)</td>
</tr>
<tr>
<td>Male</td>
<td>0.63 (0.47-0.85)</td>
</tr>
<tr>
<td>BMI</td>
<td>1.09 (1.05-1.13)</td>
</tr>
<tr>
<td>Baseline HAQ-DI score of 0.1</td>
<td>NA</td>
</tr>
<tr>
<td>Exercise, min/wk</td>
<td>0.96 (0.91-1.00)</td>
</tr>
</tbody>
</table>

Abbreviations: BMI, body mass index (calculated as weight in kilograms divided by height in meters squared); HAQ-DI, Health Assessment Questionnaire Disability Index; NA, not applicable; NS, not significant.

Analysis was restricted to participants with baseline HAQ-DI scores of 0. Variables included in the final model include group (runners vs controls), age (year), sex, BMI (measured 1 year prior to disability measure), and vigorous exercise (measured 1 year prior to disability measure).

Analysis includes all participants at study inception. The final model includes group, age (year), sex, and baseline HAQ-DI score of 0.1. Smoking, BMI, and exercise did not meet statistical significance to be included in the final model.

By the end of 2003, 81 (15%) of the runners and 144 (34%) of controls had died. The Kaplan-Meier plot of survival estimates for each group (Figure 4) shows that runners had a lower hazard rate compared with controls (P < .001 by log rank test).
ners had a significant reduction in early mortality that was maintained or increased over the study period ($P /H_{11021} .001$).

Multivariate Cox proportional hazard models generated to adjust for other variables at baseline that were associated with survival (Table 3) found that runners continued to demonstrate a significant survival advantage (HR, 0.61; 95% CI, 0.45-0.82). As expected, older age (HR, 1.12; 95% CI, 1.10-1.14), male sex (HR, 1.52; 95% CI, 1.12-2.07), and initial HAQ-DI level (HR, 1.16; 95% CI, 1.07-1.25) were associated with an increased risk of mortality. Body mass index, smoking, and baseline exercise did not meet sufficient significance to be included in the final model.

Causes of death are summarized in Table 4. A total of 225 deaths (23% of all study participants) were seen over 17 201 person-years of observation. Rates of death were increased in controls compared with runners, not only for cardiovascular outcomes as anticipated but also for nearly all identified causes.

**COMMENT**

This study demonstrates that participation in long-term running and other vigorous exercise among older adults is associated with less disability and lower mortality over 2 decades of follow-up. We prospectively followed a cohort of healthy adults from a mean age of 59 years in 1984 to 78 years in 2005. Not only were mean disability levels lower among runners at all time points, but the rate of disability progression strongly favored runners throughout the study. Results were similar when all participants or completers were analyzed. At baseline, both groups had negligible disability; however, after 21 years, runners had a mean HAQ-DI score of nearly 0.2, equivalent to having mild functional disability in 1 to 2 of the 8 areas of daily activity. In contrast, the mean HAQ-DI score of controls at 21 years approached 0.5, equivalent to moderate functional disability in 2 of the 8 areas or complete inability to perform in at least 1 area of daily functioning. Although these levels are lower than what is seen in subjects with chronic musculoskeletal diseases (the mean HAQ-DI score in rheumatoid arthritis is 0.7 to 1.022,23 and osteoarthritis 0.824), the higher levels among controls translate into important differences in overall daily functional limitations.25

In addition to confirming an overall survival advantage and reduction in cardiovascular-related deaths among persons who participate in regular exercise, we also found a reduced rate of deaths from other causes including malignant neoplasms and neurologic disorders. This is consistent with reports associating regular exercise with reduced incidence of dementia2 and several cancer types.26-28 Potential reasons for improved functional status and survival among regular exercisers may include increased cardiovascular fitness and improved aerobic capacity and organ reserve,29-31 increases in skeletal mass and metabolic adaptations of muscle with decreased frailty,20-31 lower levels of circulating inflammatory markers,32 improved response to vaccinations,33 and improved higher-order cognitive functions.34

This study follows our report of disability and mortality in this cohort after 13 years, showing significantly better outcomes for runners compared with controls.11 We had anticipated that, with an additional 8 years of observation encompassing an additional 132 deaths among participants (93 deaths were reported after 13 years), we would

**Table 4. Causes of Death Since Study Inception (1984)**

<table>
<thead>
<tr>
<th>Cause of Death</th>
<th>Total No. of Deaths</th>
<th>Runners Club Members</th>
<th>Community Controls</th>
<th>Rate Ratio, Controls/Runners</th>
<th>$P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>225</td>
<td>81</td>
<td>144</td>
<td>1999</td>
<td>2.5</td>
</tr>
<tr>
<td>Cardiovascular</td>
<td>72</td>
<td>29</td>
<td>43</td>
<td>597</td>
<td>2.1</td>
</tr>
<tr>
<td>Coronary artery disease/MI</td>
<td>38</td>
<td>14</td>
<td>25</td>
<td>347</td>
<td>2.5</td>
</tr>
<tr>
<td>Stroke</td>
<td>10</td>
<td>3</td>
<td>7</td>
<td>97</td>
<td>3.2</td>
</tr>
<tr>
<td>Congestive heart failure</td>
<td>4</td>
<td>2</td>
<td>2</td>
<td>28</td>
<td>1.4</td>
</tr>
<tr>
<td>Cancer</td>
<td>71</td>
<td>30</td>
<td>41</td>
<td>569</td>
<td>1.9</td>
</tr>
<tr>
<td>Prostate</td>
<td>7</td>
<td>4</td>
<td>40</td>
<td>3</td>
<td>1.0</td>
</tr>
<tr>
<td>Lung</td>
<td>14</td>
<td>5</td>
<td>9</td>
<td>125</td>
<td>2.5</td>
</tr>
<tr>
<td>Colon</td>
<td>10</td>
<td>4</td>
<td>4</td>
<td>6</td>
<td>2.1</td>
</tr>
<tr>
<td>Breast</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>42</td>
<td>4.2</td>
</tr>
<tr>
<td>Hematologic</td>
<td>11</td>
<td>6</td>
<td>5</td>
<td>69</td>
<td>1.2</td>
</tr>
<tr>
<td>Esophageal</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>28</td>
<td>2.8</td>
</tr>
<tr>
<td>Pancreas</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>28</td>
<td>2.8</td>
</tr>
<tr>
<td>Other</td>
<td>19</td>
<td>8</td>
<td>11</td>
<td>153</td>
<td>2.0</td>
</tr>
<tr>
<td>Neurological</td>
<td>20</td>
<td>6</td>
<td>14</td>
<td>194</td>
<td>3.2</td>
</tr>
<tr>
<td>Infections</td>
<td>16</td>
<td>1</td>
<td>10</td>
<td>15</td>
<td>20.8</td>
</tr>
<tr>
<td>Pneumonia</td>
<td>9</td>
<td>0</td>
<td>9</td>
<td>125</td>
<td>NA</td>
</tr>
<tr>
<td>Other</td>
<td>39</td>
<td>11</td>
<td>28</td>
<td>389</td>
<td>3.5</td>
</tr>
<tr>
<td>Unknown</td>
<td>7</td>
<td>4</td>
<td>4</td>
<td>40</td>
<td>1.0</td>
</tr>
</tbody>
</table>

Abbreviations: MI, myocardial infarction; NA, not applicable.

a Expressed per 100 000 person-years.

b Male subjects only.
c Female subjects only.
begin to see a convergence of disability and survival curves; however, this was not the case. Differences between runners and controls for all outcomes continued to diverge after 21 years of follow-up. Interestingly, in our analysis of 21 years of data, aerobic exercise was no longer a statistically significant independent predictor of mortality. Sixty percent of deaths occurred during the 8-year period between our last report and the present analysis, and it is possible that with this additional mortality data, vigorous exercise has become more collinear with running and no longer is identified as an independent predictor of death. Further observation of this cohort, as the remaining 440 participants reach the biological limits of life expectancy, may be required to further clarify the independent role of nonrunning, vigorous exercise.

There are limitations in this study that are important to consider. Self-selection bias is always a concern in observational study designs lacking random assignment of external interventions. Unmeasured lifestyle variables, including eating habits and use of routine preventative medical care, may have influenced results. However, great care was taken to minimize possible selection biases. The control group was selected from a larger group of relatively healthy adults who worked in a university community in 1972 and were socioeconomically similar to the runners club members. Analyses included statistical adjustments made for the following covariates that differed between the groups at study onset: age, sex, BMI, smoking, weekly exercise, and initial disability.

To control for potential self-selection bias based on group membership rather than running status, we conservatively created groups of ever and never runners to account for people who discontinued running prior to study inception due to injury, disability, or other reasons. Results of analyses comparing these groups did not differ appreciably from those obtained by analysis of groups based on club membership.

Dropout rates are always a concern in longitudinal studies. Overall, 60% of initial study participants who were still alive at the 21-year assessment continued to participate (62% of surviving runners compared with 54% of surviving controls). Given that observation spanned 2 decades, the proportion who continued to participate and were not lost to follow-up is good, and the absolute rate of discontinuation is similar in both groups. However, there were greater differences in baseline characteristics between completers and noncompleters among control subjects. Control completers tended to be younger, have lower initial HAQ-DI scores, and exercise more than the runners who died or withdrew from study participation. Analyses of mortality closely mirrored those seen for disability, with a clear survival advantage associated with physical exercise remaining after adjustment for relevant confounders.

With the exception of death, outcome variables in the study were obtained by self-report. The HAQ-DI has been validated in multiple independent studies using varied populations of healthy subjects as well as those with arthritis and other chronic conditions with excellent reliability.\textsuperscript{16,17} Validation of self-report variables against observed performance at cohort inception demonstrated excellent correlation for both groups.\textsuperscript{13}

Because our cohorts of runners and controls were relatively homogeneous at baseline (the majority were white, completed college education, had a BMI within normal limits, and had low alcohol and tobacco consumption), it is possible that these results may not be generalizable to a broader range of persons with different ethnic backgrounds, educational opportunities, access to preventive health care, or lifestyle habits. Alternatively, a study of “privileged” persons has the advantage of lessening concern over poverty, insurance status, education, access to health care, and other social variables that might otherwise be different across groups and constitute a significant bias.

Our findings of decreased disability in addition to prolonged survival among middle-aged and older adults participating in routine physical activities further support recommendations to encourage moderate to vigorous physical activity at all ages. Increasing healthy lifestyle behaviors may not only improve length and quality of life but also hopefully lead to reduced health care expenditures associated with disability and chronic diseases.\textsuperscript{33}

This study was originally designed as a test of the compression of morbidity hypothesis\textsuperscript{34,35} with the assumption that runners compared with controls would show a greater compression of disability in the remaining years of life. The present results are suggestive of such an effect. To date, only a quarter of participants have died, and although it remains possible that these trends may be different in the final three-fourths of ultimate decedents, it is unlikely, given that mortality rates between groups are expected to converge entirely after age 100 years, and cumulative lifetime disability seems unlikely to ultimately favor controls given the huge differences favoring runners through the age of 80 years, with relatively few years of life remaining. We believe, therefore, that this study contributes to arguments supporting the compression of morbidity hypothesis.

Accepted for Publication: February 18, 2008.
Correspondence: Eliza F. Chakravarty, MD, MS, Division of Immunology and Rheumatology, Stanford University School of Medicine, 1000 Welch Rd, Ste 203, Palo Alto, CA 94304 (rchakravarty@stanford.edu).

Author Contributions: Study concept and design: Chakravarty, Hubert, and Fries. Acquisition of data: Hubert, Lingala, and Fries. Analysis and interpretation of data: Chakravarty, Hubert, Lingala, and Fries. Drafting of the manuscript: Chakravarty. Critical revision of the manuscript for important intellectual content: Hubert, Lingala, and Fries. Statistical analysis: Hubert and Lingala. Obtained funding: Fries. Study supervision: Fries.

Financial Disclosure: None reported.
Funding/Support: This study was supported by grant AR43584 from the National Institute of Arthritis and Musculoskeletal and Skin Diseases and the National Institute on Aging, National Institutes of Health, Bethesda, Maryland.

REFERENCES