Exercise Effects on Bone Mineral Density, Falls, Coronary Risk Factors, and Health Care Costs in Older Women

The Randomized Controlled Senior Fitness and Prevention (SEFIP) Study

Wolfgang Kemmler, PhD; Simon von Stengel, PhD; Klaus Engelke, PhD; Lothar Haberle, PhD; Willi A. Kalender, PhD, MD

Background: Physical exercise affects many risk factors and diseases and therefore can play a vital role in general disease prevention and treatment of elderly individuals and may reduce costs. We sought to determine whether a single exercise program affects fracture risk (bone mineral density [BMD] and falls), coronary heart disease (CHD) risk factors, and health care costs in community-dwelling elderly women.

Methods: We conducted a randomized, single-blinded, controlled trial from May 1, 2005, through July 31, 2008, recruiting women 65 years or older who were living independently in the area of Erlangen-Nuremberg, Germany. In all, 246 women were randomly assigned to an 18-month exercise program (exercise group) or a well-being program (control group). The exercise group (n=123) performed a multipurpose exercise program with special emphasis on exercise intensity; the controls (n=123) focused on well-being with a low-intensity, low-frequency program. The main outcome measures were BMD, the number of falls, the Framingham-based 10-year CHD risk, and direct health care costs.

Results: For the 227 women who completed the 18-month study, significant exercise effects were observed for BMD of the lumbar spine (mean [95% confidence interval (CI)] percentage of change in BMD [baseline to follow-up] for the exercise group: 1.77% [1.26% to 2.28%] vs controls: 0.33% [-0.24% to 0.91%]; P < .001), femoral neck (exercise group: 1.01% [0.37% to 1.65%] vs controls: -1.05% [-1.70% to -0.40%]; P < .001), and fall rate per person during 18 months (exercise group: 1.00 [0.76 to 1.24] vs controls: 1.66 [1.33 to 1.99]; P = .002). The 10-year CHD risk was significantly affected in both subgroups (absolute change for the exercise group: -1.96% [95% CI, -2.69% to -1.23%] vs controls: -1.15% [-1.69% to -0.62%]; P = .22), with no significant difference between the groups. The direct health care costs per participant during the 18-month intervention showed nonsignificant differences between the groups (exercise group: €2255 [95% CI, €1791-€2718] vs controls: €2780 [€2187-€3372]; P = .20).

Conclusion: Compared with a general wellness program, our 18-month exercise program significantly improved BMD and fall risk, but not predicted CHD risk, in elderly women. This benefit occurred at no increase in direct costs.

Trial Registration: clinicaltrials.gov Identifier: NCT00267839

Arch Intern Med. 2010;170(2):179-185

See also pages 124, 170, 186, and 194

Author Affiliations: Institute of Medical Physics (Drs Kemmler, von Stengel, Engelke, and Kalender) and Institute of Biometry and Epidemiology (Dr Haberle), Friedrich-Alexander University of Erlangen-Nuremberg, Erlangen, Germany.
A semi-blinded, randomized 18-month controlled trial was conducted to compare the effects of an intense exercise program with a low-intensity, low-frequency control program on multiple risk factors and HCCs of elderly community-dwelling women. The study was blinded for the outcome assessors and participants but not for the investigator (W.K.). Unlike other exercise studies, participants were blinded to the underlying hypothesis by postulating a different study aim (wellness) for the control group. In this study, the primary goal of the wellness protocol was to positively affect well-being and to briefly introduce topics related to a healthier lifestyle and physical activity. Both groups were trained separately to prevent contact between the cohorts. The effectiveness of the blinding in the control group was proven in structured interviews conducted by the primary investigators at the end of the 18 months.

The study protocol was approved by the ethics committee of the Friedrich-Alexander University of Erlangen-Nuremberg and the Bundesamt für Strahlenschutz. The study was designed and executed by the Institute of Medical Physics, and randomization and statistical procedures were performed by the Institute of Biometry and Epidemiology, both at Friedrich-Alexander University of Erlangen-Nuremberg. All study participants gave written informed consent. The study was fully registered under http://www.clinicaltrials.gov.

In the present study, we tested the hypothesis that our multipurpose exercise protocol significantly affects BMD, fall frequency, CHD risk, and HCCs compared with a low-volume, low-intensity “wellness” program.

### PARTICIPANTS

Six hundred fifty-nine women, all members of Siemens Health Insurance, who were 65 years or older and were living independently in the area of Erlangen-Nuremberg replied to our individualized information letter (n = 7500). Participants were recruited by mail between May 1, 2005, and January 31, 2006 (±1 week). Two hundred eighty-three women were not admitted to the study because of diseases (eg, alcoholism and hypercortisolism) or medication usage (eg, bisphosphonates, hormone therapy, glucocorticoids, and laxatives) affecting bone metabolism or fall risk during the past 2 years (n = 254), a history of cardiovascular disease (eg, stroke or cardiac events) (n = 13), acute or chronic inflammatory diseases (n = 2), known secondary osteoporosis (n = 2), participation in exercise studies during the preceding 2 years (n = 10), or very low physical capacity (<50 W at cycle ergometry) or athletic history during the past decade (n = 2). Three hundred seventy-six women were invited to meetings to receive detailed study information. Primarily because of the randomization procedure and the inability to choose the preferred study arm, 80 individuals refused to participate. Thus, 296 women were assigned by computer-generated block randomization (first sequence: 130 subjects and a block size of 3 × 2; last sequence: 146 subjects and a block size of 2 × 2) stratified for age to 3 intervention arms: an exercise program (exercise group; 123 women), a control group (123 women), or an exercise and whole-body vibration group (50 women). The whole-body vibration group constituted a substudy with extended eligibility criteria (ie, exclusion in case of knee or hip implant). Owing to the different eligibility criteria, divergent study aims, and additional vibration protocol, the whole-body vibration group was not included in the present analysis.

All participants remained in the same group throughout the study. Table 1 gives the baseline characteristics of the exercise and control groups, and the Figure shows the participant flow through the study.

### METHODS

### Table 1. Baseline Characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Exercise Group (n = 123)</th>
<th>Control Group (n = 123)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, y</td>
<td>68.9 (3.9)</td>
<td>69.2 (4.1)</td>
<td>.45</td>
</tr>
<tr>
<td>Median (95% CI)</td>
<td>68 (65-78)</td>
<td>68 (65-78)</td>
<td></td>
</tr>
<tr>
<td>Height, cm</td>
<td>161.8 (6.1)</td>
<td>160.5 (5.8)</td>
<td>.08</td>
</tr>
<tr>
<td>Weight, kg</td>
<td>68.1 (10.9)</td>
<td>69.5 (12.0)</td>
<td>.33</td>
</tr>
<tr>
<td>Total body fat, %</td>
<td>36.3 (5.9)</td>
<td>37.4 (5.6)</td>
<td>.09</td>
</tr>
<tr>
<td>Age at menopause, y</td>
<td>48.8 (5.9)</td>
<td>49.8 (5.3)</td>
<td>.16</td>
</tr>
<tr>
<td>Total body fat, %</td>
<td>1574 (411)</td>
<td>1535 (392)</td>
<td>.51</td>
</tr>
<tr>
<td>Calcium intake, mg/d</td>
<td>828 (414)</td>
<td>816 (356)</td>
<td>.89</td>
</tr>
<tr>
<td>Baseline physical activity indexc</td>
<td>4.2 (1.1)</td>
<td>4.3 (1.2)</td>
<td>.44</td>
</tr>
<tr>
<td>Exercise volume, min/km²</td>
<td>85 (102)</td>
<td>67 (111)</td>
<td>.45</td>
</tr>
<tr>
<td>VO₂ peak, mL/kg/min²</td>
<td>24.1 (4.1)</td>
<td>22.9 (4.2)</td>
<td>.06</td>
</tr>
<tr>
<td>Smokers, No.</td>
<td>4</td>
<td>4</td>
<td>&gt;.99</td>
</tr>
<tr>
<td>Diabetes mellitus, No.</td>
<td>10</td>
<td>11</td>
<td>.76</td>
</tr>
<tr>
<td>Hypertension, No.</td>
<td>51</td>
<td>60</td>
<td>.25</td>
</tr>
<tr>
<td>High cholesterol, No.</td>
<td>64</td>
<td>52</td>
<td>.80</td>
</tr>
<tr>
<td>Metabolic syndrome, No.</td>
<td>45</td>
<td>50</td>
<td>.43</td>
</tr>
<tr>
<td>Osteoporosis, No.</td>
<td>27</td>
<td>23</td>
<td>.53</td>
</tr>
<tr>
<td>Osteoarthritis, No.</td>
<td>9</td>
<td>9</td>
<td>&gt;.99</td>
</tr>
<tr>
<td>Multiple morbidity, No.</td>
<td>67</td>
<td>75</td>
<td>.30</td>
</tr>
</tbody>
</table>

Abbreviations: CI, confidence interval; VO₂ peak, peak oxygen consumption.

* SI conversion factor: To convert energy intake from kilocalories to kilojoules, multiply by 4.186.
* As assessed by baseline questionnaire.
* Four-day dietary protocol.
* Self-rated physical activity score (1 indicates very low; 7, very high).
* Stepwise treadmill test to voluntary maximum.
* Defined as diastolic blood pressure of greater than 90 mm Hg and/or systolic blood pressure of greater than 140 mm Hg.
* Defined as total cholesterol of greater than 250 mg/dL (to convert to millimoles per liter, multiply by 0.0259).
* Defined as having a bone mineral density T-score of less than −2.5 at the lumbar spine or total hip.
* Defined as ≥ 2 or more diseases. Besides the factors given in the table, all other diseases (including diabetes and osteoarthritis) were assessed by baseline questionnaire.
A high-intensity exercise program (primary intervention) was compared with an intermittent low-frequency wellness program that focused on well-being. The study intervention took place from January 1, 2006, to September 30, 2007 (±1 week). Participants of both groups were supervised by certified trainers who were instructed weekly by the principal investigator (W.K.). Individual training logs were analyzed every 10 weeks to determine participant adherence to the exercise or wellness program and attendance. For ethical reasons and because BMD was a primary study end point, calcium (1500 mg/d) and cholecalciferol (500 IU/d) supplements (Opfermann Arzneimittel, Wiehl, Germany) were provided for all participants in both groups. However, adherence to the calcium and cholecalciferol intake regimen was not determined in this study. Apart from the exercise intervention, subjects were requested to maintain their usual lifestyle and exercise habits.

**EXERCISE PROGRAM**

The weekly exercise program consisted of two 60-minute supervised group classes and two 20-minute home training sessions. Group classes were structured into 4 sequences:

1. A warm-up/aerobic dance sequence with progressively higher-impact elements performed for 20 minutes at 70% to 85% maximum heart rate, as assessed during a treadmill test to a voluntary maximum.

2. Five minutes of static and dynamic balance training. Briefly, exercises were performed in a standing position under conditions of progressively increasing postural instability.

3. Functional gymnastics, isometric strength training, and stretching sequences with 1 to 3 sets of isometric floor exercises for trunk flexors and extensors, hip flexors and extensors, and leg abductors and adductors. Typically, 10 to 15 exercises were performed with 6 to 10 seconds of maximum intensity and 20 to 30 seconds of active rest. During the rest periods, continuous stretching for the corresponding muscle groups was performed. Exercises were replaced every 6 to 18 weeks by more strenuous ones.

4. Two to 3 sets of upper body exercises (low and high belt rowing and belt shoulder raises) with 10-15 repetitions and a 2-second concentric–1-second static–2-second eccentric time under tension intermittent with 30 to 40 seconds of continuous stretching were executed using elastic belts (Thera-Band; Ludwig Artzt GmbH, Hadamar, Germany). Participants were encouraged to perform at their maximum exertion level minus 2 repetitions. To ensure a progressive development of strength, different belts (3.5 kg, 4.5 kg, and 6 kg per 100% extension) were used. Furthermore, 2 sets with 3 unilateral dynamic weight-bearing leg exercises were performed in a circuit mode with 1 minute of exercise and 1 minute of active rest, with stretching exercises for the corresponding muscle groups. Participants were asked to execute 8 repetitions per leg (2-second–0-second–2-second time under tension mode) on a maximum exertion level minus 2 repetitions. Intensity was progressively increased by enlarging the amplitude of the movement, changing the velocity of the concentric execution, and introducing more strenuous exercises.

The home training session emphasized strength and flexibility exercises, with 1 to 2 sets of 6 to 8 isometric exercises, 2 to 3 belt exercises with 2 sets of 10 to 15 repetitions, and intermittent stretching exercises. Every 12 weeks, a new home
training routine with more intense exercises replaced the existing protocol. Participants were regularly encouraged by the instructors to consistently practice the home training session.

**WELLNESS PROGRAM**

To blind the participants, the control group performed a program that focused on well-being and was designed not to cause physical adaptations. These participants executed a low-frequency, low-intensity protocol for 60 minutes once a week for 10 weeks followed by 10 weeks of rest. This training-rest cycle was repeated throughout the study. Each training session started with 5 to 10 minutes of walking at 50% to 60% maximum heart rate and finished with 10 minutes of a muscular relaxation sequence. The main topic changed from week to week. Within each of the four 10-week blocks, the following activities were performed: relaxation, games/interaction, general coordination, endurance, balance, dances, body sensitivity, muscle strength, breathing, and flexibility.

Endurance and strength-training sessions were executed with low to moderate intensity without progressive increments of exercise intensity or duration.

**MEASUREMENTS**

Our primary outcome measures consisted of BMD at the lumbar spine (LS) and proximal femur, fall frequency (fall rate), projected 10-year CHD risk, and HCCs.

Our secondary outcome measures included the number of fallers per group, number of fallers with injurious falls, and overall number of fractures per group. In addition, the following variables constituted the 10-year CHD risk score and were analyzed: low- and high-density lipoprotein cholesterol levels, blood pressure, smoking, and the presence of diabetes mellitus.

Baseline and 18-month measurements were obtained by the same researcher and at the same time of day (±1 hour). All assessments were determined in a blinded fashion.

Height was determined with a stadiometer, and weight was measured on a digital scale while the participant wore minimal clothing. Body fat and BMD at the LS and proximal femur were assessed with dual-energy x-ray absorptiometry (QDR 4500, Discovery upgrade; Hologic, Bedford, Massachusetts) using standard protocols specified by the manufacturer. The coefficient of variance for the dual-energy x-ray absorptiometry scan of the LS was 0.9%, and the corresponding value for the femoral neck was 1.0%.

Falls were defined according to the PROFANE (Prevention of Falls Network Europe) group. Injurious falls and overall fractures were monitored daily with the use of fall calendars compiled by the participants. Outcome assessors contacted subjects who fell and nonresponders monthly by telephone. The 6-month prestudy fall rate was recorded at the study start in response to a questionnaire.

The 10-year CHD risk was calculated according to the Framingham Risk Calculator by Wilson et al. Briefly, points were added to each risk factor for sex, age, low- and high-density lipoprotein cholesterol level, systolic and diastolic blood pressure, diabetes mellitus, and smoking. Based on a score sheet, the 10-year CHD risk for each subject was given as a percentage value.

In close collaboration with Siemens Health Insurance (Siemens Betriebs Krankenkasse), total HCCs (excluding dental costs) were determined for the 6 months before the study and for the 18-month intervention. Final data for HCCs were compiled by the participants. Outcome assessors contacted subjects as well as fallers and fall history; and prestudy physical activity and exercise levels. To control changes caused by confounding factors (eg, medication, physical activity, and exercise level) and adverse effects (defined as a harmful or abnormal result), follow-up questionnaires and structured interviews were performed by the same physician.

Nutritional intake was assessed using 4-day dietary protocols and analyzed with the use of Prodi-4.5/03 Expert software (Wissenschaftlicher Verlag, Freiburg, Germany). Based on the results, a maximum of 1500 mg of calcium and 500 IU of cholecalciferol were given to supplement the diets of all participants.

**STATISTICAL ANALYSES**

The estimated sample size was based on fall frequency. To detect a 33% difference between groups in the fall rate during the 18-month intervention (effect size=0.33), 115 participants per group were required for a 5% error probability with 80% statistical power (1-tailed). To adjust for subjects who were lost to follow-up, 123 participants were included in each group. We performed an intention-to-treat analysis that included all subjects with 18-month follow-up data.

The allocation sequence and group assignment were performed by the Institute of Biometry and Epidemiology. Participants were enrolled by the Institute of Medical Physics.

Bone mineral density was log transformed to obtain the normally distributed data required for the analysis of variance with repeated measurements. This was performed as a mixed linear model, with the treatment group and time as fixed factors and with subjects as the random factor. By using a mixed linear model, the baseline values of participants who were lost to follow-up could be included in the analysis. We report herein the P values of post hoc tests according to the Tukey-Kramer method.

The number of falls (fall rate with 95% confidence intervals [CIs]) was compared between groups by means of negative binomial regression.

The 10-year CHD risk and health cost data could not be transformed to normally distributed data and were therefore analyzed nonparametrically with the Wilcoxon rank sum test in the 2-sample cases and with the Wilcoxon signed rank test in the 1-sample cases. Contrary to the analysis of variance with repeated measurements, only complete data (ie, data from baseline and follow-up), are evaluated with the nonparametric tests.

Effect sizes were based on the absolute difference (SD) between the baseline and 18-month follow-up values in the exercise and control groups and were calculated using Cohen's d. All tests were 2-sided, with P<.05 considered statistically significant. Raw P values without further adjusting are presented. All statistical evaluations were performed with SAS software (version 9.1; SAS Institute Inc, Cary, North Carolina).

**RESULTS**

The Figure shows the participant flow during the SEFIP study. All participants are included in the final analysis; 115 women in the exercise group (93.5%) and 112 participants in the controls (91.1%) provided follow-up outcomes data. Twelve participants lost interest; 8 of them cited study-related reasons with respect to the exercise protocol (4 women in the exercise group and 3 controls) or because of the calcium and cholecalciferol supplementation (1 woman in the control group). Two participants moved to a different city and 1 woman died. Four
participants were unable to attend the final assessment. The mean (SD) overall attendance rate in the exercise group was 76.3% (8.1%) for the group training session (for the controls, 72.0% [8.7%]) and 42.2% (5.3%) for the home training session. Furthermore, as determined by training logs kept by both groups, adherence to the exercise execution, strain thresholds, modalities, and progression) and wellness protocols was excellent for completion of the sessions performed.

Other than the intervention, no changes in physical activity or exercise level were determined for either group. No musculoskeletal injuries, falls, or CHD events occurred during the training sessions. In addition, no adverse effects (ie, pain or quality-of-life reductions) of our exercise or wellness protocol were observed.

Blinding of the control group was successful; 81.3% of the participants believed they were part of the primary intervention. In the exercise group, all of the participants believed they were part of the primary intervention.

Table 2 shows between-group differences and the numerical changes of the end points. At baseline, none of the differences were significant. Mean (95% CI) BMD at the LS (1.77% [1.26% to 2.28%]) and femoral neck (1.01% [0.37% to 1.65%]) significantly increased in the exercise group (P < .001 for both), whereas, among the controls, BMD did not significantly change at the LS (0.33% [−0.24% to 0.91%]; P = .25) and significantly decreased at the femoral neck (−1.05% [−1.70% to −0.40%]; P < .001). Significant group differences (P < .001) were determined for both measures (Table 2).

Prestudy fall rates (falls per person per group for a 6-month period) did not differ between groups (Table 2). However, significant differences between the exercise group (1.00 [95% CI, 0.76–1.24]) and controls (1.66 [1.33–1.99]) were observed for the 18-month intervention period. Furthermore, relative risk was calculated as 0.54 (95% CI, 0.35–0.84; P = .01) for subjects who fell and 0.33 (0.15–0.74; P = .01) for those who experienced injurious falls in the exercise group. Overall fractures due to falls were twice as high in the controls (n = 12) than in the exercise group (n = 6). There were no fall-induced spine or proximal femur fractures, and the group difference was not significant (P = .90).

The 10-year CHD risk significantly (P < .001) improved in both groups (exercise group: −1.96% [95% CI, −2.69% to −1.23%] vs controls: −1.15% [−1.69% to −0.62%]; P = .22). Changes in 10-year CHD risk in the exercise group were primarily based on significant changes in high-density lipoprotein cholesterol level (exercise group: −1.96% [95% CI, 4.3% to 8.7%] vs controls: 1.8% [3.1% to 6.0%]; P = .002), and low-density lipoprotein cholesterol level (exercise group: −1.9% [−4.5% to 0.7%] vs exercise group: 3.1% [−0.1% to 6.3%]; P = .02). Both groups further experienced significant changes in systolic exercise group: −3.5% [95% CI, −5.8% to −1.3%] vs controls: −4.8% [−7.1% to −2.5%]; P = .43) and diastolic exercise group: −8.7% [−10.9% to −6.6%] vs controls: −7.6% [−9.9% to −5.3%]; P = .48) blood pressure that were comparable between the groups.

Other variables (diabetes and smoking) contributing to the 10-year CHD risk score were not significantly affected.

Prestudy HCCs did not differ between the groups (Table 2). At 18 months, HCCs per participant were higher in the control group (exercise group: €2255 [95% CI, €1791–€2718] vs controls: €2780 [€2187–€3372]; P = .20 [not included were dental costs, which were comparable between groups]), but between-group differences were not significant. These costs included the intervention costs per participant of €383 (€328–€437) in the exercise group and €175 (€134–€216) in the controls.
Significant changes in the confounding variables (lifestyle and nutritional intake) were not observed during the intervention period. However, 8 women in the exercise group and 10 in the controls modified their blood pressure, cholesterol level, and dosage of hypothyroidism or hyperthyroidism medication.

The results of our exercise program clearly demonstrated positive effects on the most relevant risk factors (fracture and CHD risk factors) for elderly women, although the CHD risk factors were not significantly better compared with those of the control group. The 50%-lower fracture incidence in the exercise group was probably a result of the decreased fall rate in conjunction with increased BMD. However, this surprising result was not significant, and fracture rate was not a main study end point. Even the control group, who participated in the wellness intervention, significantly reduced their 10-year CHD risk, mainly through a favorable effect on blood pressure. Thus, contrary to the high-impact, high-intensity strategy required to affect BMD, the rather smooth effort exerted in the wellness program was apparently able to significantly affect blood pressure in this cohort.

We performed an exercise program with fewer group sessions than typically recommended but with high exercise intensity during the aerobic and strength section. Although this strategy was successful for our clinical end points, one may worry whether this approach increases joint or low back pain in elderly subjects and may therefore be inadequate in this population. However, similar to our recent high-intensity exercise program for early postmenopausal women, any negative changes in pain indicators or quality of life were detected in the present study. This result was indirectly supported by a low dropout rate and a moderate to high attendance rate.

Our study possesses several strengths. Our cohort was a homogeneous community-dwelling group of women 65 years or older. The blinding strategy was successful. The study duration was sufficiently long to detect relevant changes of physiological variables. Potential effects of lifestyle changes, diseases, medication, and nutrition were strictly controlled. The exercise regimen was progressively augmented during the intervention period, and the group sessions were strictly supervised by certified trainers. The favorable attendance at the group sessions and the favorable dropout rates indicate the attractiveness of the exercise program. However, the low adherence to the home training reported by the exercise group participants again demonstrates the reluctance of most elderly individuals to exercise on their own. Health care costs were directly assessed rather than estimated. The funding organizations played no role in the design or conduct of study. Siemens Be-
triebs Krankenkasse was responsible for the collection of the total HCCs; however, none of the funding partners influenced the analyses, the interpretation of the data, or the preparation, review, or approval of the manuscript.

Additional Contributions: Jerry Mayhew, PhD, of Truman State University, Kirksville, Missouri, provided helpful critical remarks.

REFERENCES


